

Quantum Computer as a Probabilistic Inference Engine

Robert R. Tucci
P.O. Box 226
Bedford, MA 01730
tucci@ar-tiste.com

February 1, 2008

Abstract

We propose a new class of quantum computing algorithms which generalize many standard ones. The goal of our algorithms is to estimate probability distributions. Such estimates are useful in, for example, applications of Decision Theory and Artificial Intelligence, where inferences are made based on uncertain knowledge. The class of algorithms that we propose is based on a construction method that generalizes a Fredkin-Toffoli (F-T) construction method used in the field of classical reversible computing. F-T showed how, given any binary deterministic circuit, one can construct another binary deterministic circuit which does the same calculations in a reversible manner. We show how, given any classical stochastic network (classical Bayesian net), one can construct a quantum network (quantum Bayesian net). By running this quantum Bayesian net on a quantum computer, one can calculate any conditional probability that one would be interested in calculating for the original classical Bayesian net. Thus, we generalize the F-T construction method so that it can be applied to any classical stochastic circuit, not just binary deterministic ones. We also show that, in certain situations, our class of algorithms can be combined with Grover's algorithm to great advantage.

1 Introduction

In this paper, we use the language of classical Bayesian (CB) and quantum Bayesian (QB) nets[1]. The reader is expected to possess a rudimentary command of this language.

We begin this paper with a review of various standard quantum computing algorithms; namely, those due to Deutsch-Jozsa[2], Simon[3], Bernstein-Vazirani[4], and Grover[5]. We discuss these standard algorithms both in terms of qubit circuits (the conventional approach) and QB nets. Then we propose a class of quantum computing algorithms which generalizes the standard ones.

Most standard quantum computing algorithms are designed for calculating deterministic or almost deterministic probability distributions. (By a deterministic probability distribution we mean one whose range is restricted to either zero or unit probabilities.) In contrast, our algorithms can also estimate more general probability distributions. Such estimates are useful in, for example, applications of Decision Theory and Artificial Intelligence, where inferences are made based on uncertain knowledge.

Since some of the standard algorithms are contained in the class of algorithms that we propose, some algorithms in our class have a time-complexity advantage over the best classical algorithms for performing the same task. Even those algorithms in our class that have no complexity advantage might still be useful for nanoscale quantum computing because they are reversible and thus dissipate less power. Power dissipation is best minimized in nanoscale devices since it can lead to serious performance degradation.

The class of algorithms that we propose in this paper is based on a construction method that generalizes a Fredkin-Toffoli (F-T) construction method[6] used in the field of classical reversible computing. F-T showed in Refs.[6] how, given any binary gate f (i.e., a function $f : \{0, 1\}^r \rightarrow \{0, 1\}^s$, for some integers r, s), one can construct another binary gate \bar{f} such that \bar{f} is a deterministic reversible extension (DRE) of f . \bar{f} can be used to perform the same calculations as f , but in a reversible manner. Binary gates f and \bar{f} can be represented as binary deterministic circuits. In this paper, we show how, given any CB net \mathcal{N}^C , one can construct a QB net \mathcal{N}^Q which is a “q-embedding” (q=quantum) of \mathcal{N}^C . By running \mathcal{N}^Q on a quantum computer, one can calculate any conditional probability that one would be interested in calculating for the CB net \mathcal{N}^C . Our method for constructing a q-embedding for a CB net is a generalization of the F-T method for constructing a DRE of a binary deterministic circuit. Thus, we generalize their method so that it applies to any classical stochastic circuit, not just binary deterministic ones.

A quantum compiler [7] [8] can “compile” a unitary matrix; i.e., it can express the matrix as a SEO (sequence of elementary operations) that a quantum computer can understand. To run a QB net on a quantum computer, we need to replace the QB net by an equivalent SEO[9]. This can be done with the help of a quantum compiler.

Thus, the class of algorithms that we propose promises to be fertile ground for the use of quantum compilers.

In certain cases, the probabilities that we wish to find are too small to be measurable by running \mathcal{N}^Q on a quantum computer. However, we will show that sometimes it is possible to define a new QB net, call it $\mathcal{N}^{Q'}$, that magnifies and makes measurable the probabilities that were unmeasurable using \mathcal{N}^Q alone. We will refer to $\mathcal{N}^{Q'}$ as Grover's Microscope for \mathcal{N}^Q , because $\mathcal{N}^{Q'}$ is closely related to Grover's algorithm, and it magnifies the probabilities found with \mathcal{N}^Q .

2 Notation and Other Preliminaries

In this section, we will introduce certain notation that is used throughout the paper.

We will use the word “ditto” as follows. If we say “A (ditto, X) is smaller than B (ditto, Y)”, we mean “A is smaller than B” and “X is smaller than Y”.

Let $Bool = \{0, 1\}$. For integers a and b such that $a \leq b$, let $Z_{a,b} = \{a, a+1, a+2, \dots, b\}$.

For any statement \mathcal{S} , we define the truth function $\theta(\mathcal{S})$ to equal 1 if \mathcal{S} is true and 0 if \mathcal{S} is false. For example, $\theta(x > 0)$ represents the unit step function and $\delta(x, y) = \theta(x = y)$ the Kronecker delta function.

\oplus will denote addition mod 2. For any integer x , $x \% 2$ will mean the remainder from dividing x by 2. For example, $4 \% 2 = 0$ and $5 \% 2 = 1$. (This same $\%$ notation is used in the C programming language.) When speaking of bits with states 0 and 1, we will often use an overbar to represent the opposite state: $\bar{0} = 1$, $\bar{1} = 0$. Note that if $x, k \in Bool$ then

$$\sum_k (-1)^{kx} = 1 + (-1)^x = 2\delta(x, 0). \quad (1)$$

If $\vec{x}, \vec{y} \in Bool^n$, we will use $\vec{x} \cdot \vec{y} = \sum_{\alpha=0}^{n-1} x_\alpha y_\alpha$, where the addition is normal, not mod 2.

Given $x \in Z_{0,\infty}$, let $x = \sum_{\alpha=0}^{\infty} x_\alpha 2^\alpha$, where $x_\alpha \in Bool$ for all α . Then we will denote the binary representation of x by $bin(x) = (x_0, x_1, x_2, \dots)$. Thus, $bin_\alpha(x) = x_\alpha$.

On the other hand, given $\vec{x} = (x_0, x_1, x_2, \dots) \in Bool^\infty$, let $x = \sum_{\alpha=0}^{\infty} x_\alpha 2^\alpha$. Then we will denote the decimal representation of \vec{x} by $dec(\vec{x}) = x$.

We will use the symbol \sum_{\cdot} to denote a sum of whatever is on the right hand side of this symbol over those indices with a dot underneath them. For example, $\sum_{\cdot} f(a, b, c) = \sum_{a,b} f(a, b, c)$. Furthermore, \sum_{all} will denote a sum over all indices. If we wish to exclude a particular index from the summation, we will indicate this by a slash followed by the name of the index. For example, in $\sum_{all/a,b}$ we wish to exclude summation over a and b . Suppose f maps set S into the complex numbers. We will often use $\frac{f(x)}{\sum_x num}$ to represent $\frac{f(x)}{\sum_{x \in S} f(x)}$. Thus, *num* is shorthand for the numerator of the fraction.

We will underline random variables. $P(\underline{a} = a) = P_{\underline{a}}(a)$ will denote the probability that the random variable \underline{a} assumes value a . $P(\underline{a} = a)$ will often be abbreviated by $P(a)$ when no confusion will arise. $S_{\underline{a}}$ will denote the set of values which the random variable \underline{a} may assume, and $N_{\underline{a}}$ will denote the number of elements in $S_{\underline{a}}$. $pd(B|A)$ will stand for the set of probability distributions $P(\cdot|\cdot)$ such that $P(b|a) \geq 0$ and $\sum_{b' \in B} P(b'|a) = 1$ for all $a \in A$ and $b \in B$.

This paper will also utilize certain notation and nomenclature associated with classical and quantum Bayesian nets. For example, we will use $(x.)_A$ to denote $\{x_i : i \in A\}$. See Ref.[1] for a review of such notation.

$H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the one bit Hadamard matrix. $H_{N_B} = H_1^{\otimes N_B}$ (the n -fold tensor product of H_1) is the N_B bit Hadamard matrix. We will also use $\hat{H}_1 = \frac{1}{\sqrt{2}}H_1$ and $\hat{H}_{N_B} = \hat{H}_1^{\otimes N_B} = \frac{1}{\sqrt{2^{N_B}}}H_1^{\otimes N_B}$. Note that $(H_1)_{b,b'} = (-1)^{bb'}$ for $b, b' \in Bool$, and $(H_{N_B})_{\vec{b}, \vec{b}'} = (-1)^{\vec{b} \cdot \vec{b}'}$ for $\vec{b}, \vec{b}' \in Bool^{N_B}$.

Any 2×2 matrix M which acts on bit α will be denoted by $M(\alpha)$. (We like to use lower case Greek letters for bit labels.) In this notation, a controlled-not (cnot) gate with control bit κ and target bit τ can be expressed as $\sigma_x(\tau)^{n(\kappa)}$. See Ref.[8] for more details about this notation.

Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B bits. Assume all κ_i are distinct. We will often use $N_S = 2^{N_B}$, where N_B stands for number of bits and N_S for number of states. If $|\phi\rangle_{\kappa_i} = |\phi(\kappa_i)\rangle$ is a ket for qubit κ_i , define $|\phi\rangle_{\vec{\kappa}} = |\phi(\vec{\kappa})\rangle = \prod_{i=0}^{N_B-1} |\phi(\kappa_i)\rangle$. For example, if

$$|0\rangle_{\kappa_i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2)$$

for all i , then

$$|0\rangle_{\vec{\kappa}} = \prod_{i=0}^{N_B-1} |0\rangle_{\kappa_i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = [1, 0, 0, \dots, 0]^T. \quad (3)$$

Likewise, if $\Omega(\kappa_i)$ is an operator acting on qubit κ_i , define $\Omega(\vec{\kappa}) = \prod_{i=0}^{N_B-1} \Omega(\kappa_i)$. For example, $H_1(\vec{\kappa}) = \prod_{i=0}^{N_B-1} H_1(\kappa_i)$ is an N_B bit Hadamard matrix.

Next, we will introduce some notation related to Pauli matrices. The Pauli matrices are given by:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4)$$

If $|+_z\rangle$ and $|-_z\rangle$ represent the eigenvectors of σ_z with eigenvalues $+1$ and -1 , respectively, then we define

$$|0\rangle = |+_z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (5)$$

and

$$|1\rangle = |-_z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (6)$$

We denote the “number operator” by n . Thus

$$n = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = |-_z\rangle\langle -_z| = \frac{1 - \sigma_z}{2} , \quad (7)$$

and

$$\bar{n} = 1 - n = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |+_z\rangle\langle +_z| = \frac{1 + \sigma_z}{2} . \quad (8)$$

Since n and σ_z are diagonal, it is easy to see that

$$(-1)^n = \sigma_z . \quad (9)$$

It is also useful to introduce symbols for the projectors with respect to $|0\rangle$ and $|1\rangle$;

$$P_0^z = |0\rangle\langle 0| = \bar{n} , \quad (10)$$

$$P_1^z = |1\rangle\langle 1| = n . \quad (11)$$

Most of the definitions and results stated so far for σ_z have counterparts for σ_x and σ_y . The counterpart results can be easily proven by applying a rotation that interchanges the coordinate axes. Let $w \in \{x, y, z\}$. If $|+_w\rangle$ and $|-_w\rangle$ represent the eigenvectors of σ_w with eigenvalues $+1$ and -1 , respectively, then we define

$$|0_w\rangle = |+_w\rangle , \quad (12)$$

and

$$|1_w\rangle = |-_w\rangle . \quad (13)$$

Let

$$n_w = |-_w\rangle\langle -_w| = \frac{1 - \sigma_w}{2} , \quad (14)$$

$$\bar{n}_w = 1 - n_w = |+_w\rangle\langle +_w| = \frac{1 + \sigma_w}{2} . \quad (15)$$

As when $w = z$, one has

$$(-1)^{n_w} = \sigma_w . \quad (16)$$

Let

$$P_0^w = |0_w\rangle\langle 0_w| = \bar{n}_w , \quad (17)$$

and

$$P_1^w = |1_w\rangle\langle 1_w| = n_w . \quad (18)$$

In understanding Grover's algorithm, it is helpful to be aware of some simple properties of reflections on a plane. Suppose ϕ is a normalized ($\phi^\dagger\phi = 1$) complex vector. Define the projection and reflection operators for ϕ by

$$\Pi_\phi = \phi\phi^\dagger , \quad R_\phi = 1 - 2\Pi_\phi . \quad (19)$$

Note that $\Pi_\phi^2 = \Pi_\phi$. Fig.1 shows that if $x' = R_\phi x$, then x' is the reflection of x with respect to the plane perpendicular to ϕ . For example, $R_\phi\phi = -\phi$.

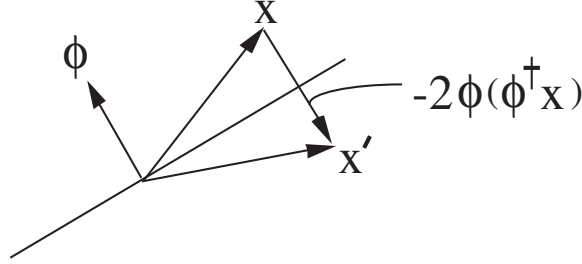


Figure 1: Reflection with respect to plane perpendicular to ϕ .

Some simple properties of R_ϕ are as follows. $R_\phi = R_\phi^\dagger$ and $R_\phi R_\phi^\dagger = R_\phi^2 = 1$. Since reflections are unitary matrices, a product of reflections is also a unitary matrix. Note that

$$(-1)^{\Pi_\phi} = e^{i\pi\Pi_\phi} = 1 + (e^{i\pi\Pi_\phi} - 1) \quad (20a)$$

$$= 1 + \Pi_\phi(e^{i\pi} - 1) = 1 - 2\Pi_\phi \quad (20b)$$

$$= R_\phi . \quad (20c)$$

(Eq.(20b) follows from the Taylor expansion of $e^{i\pi\Pi_\phi}$.)

If e_1, e_2, \dots, e_n is an orthonormal basis for a vector space, $\Pi_i = e_i e_i^\dagger$, and $R_i = 1 - 2\Pi_i$, then the product of the R_i in any order is -1 . Indeed,

$$R_1 R_2 \dots R_n = (1 - 2\Pi_1)(1 - 2\Pi_2) \dots (1 - 2\Pi_n) \quad (21a)$$

$$= 1 - 2(\Pi_1 + \Pi_2 + \dots \Pi_n) \quad (21b)$$

$$= -1 . \quad (21c)$$

Another property of reflection operators which is useful for understanding Grover's algorithm is the following. Let

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (22)$$

Now suppose that e'_1 is obtained by rotating e_1 clockwise by an angle $\theta/2$:

$$e'_1 = \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} e_1 = \begin{pmatrix} \sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix}. \quad (23)$$

$e'_1 \approx e_1$ for small θ . It is easy to check that the double reflection $-R_{e'_1}R_{e_0}$ is equivalent to a rotation (also clockwise) by θ :

$$-R_{e'_1}R_{e_0} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (24)$$

(That these two successive reflections equal a rotation was to be expected, since the reflections are orthogonal matrices and a product of orthogonal matrices is itself orthogonal.)

Above, we have considered plane reflections R_ϕ acting on a complex vector space, but our formulas still hold true when R_ϕ acts on a real instead of a complex vector space. In the case of real vector spaces, the Hermitian conjugate symbol \dagger is replaced by the matrix transpose symbol T , and unitary matrices are replaced by orthogonal matrices.

3 Some Standard Algorithms

Next we will discuss several standard algorithms that are considered among the best that the quantum computation field has to offer at the present time. Later, we will try to generalize these standard algorithms.

3.1 Deutsch-Jozsa Algorithm

In this section we will discuss the D-J (Deutsch-Jozsa) algorithm[2]. We will do this first in terms of qubit circuits (the conventional approach), and then in terms of QB nets.

Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B "control" bits and let τ label a single "target" bit. Assume that τ and all the κ_i are distinct. We will denote the state of these bits in the preferred basis (the eigenvectors of σ_z) by $|x\rangle_{\vec{\kappa}}|y\rangle_\tau$, where $x \in Bool^{N_B}$ and $y \in Bool$. Given a function $f : Bool^{N_B} \rightarrow Bool$, define the unitary operator Ω by

$$\Omega = \sigma_x(\tau) \hat{H}_1(\tau) \hat{H}_1(\vec{\kappa}) \sigma_x^{f(\vec{\kappa})}(\tau) \hat{H}_1(\vec{\kappa}) \hat{H}_1(\tau) \sigma_x(\tau), \quad (25)$$

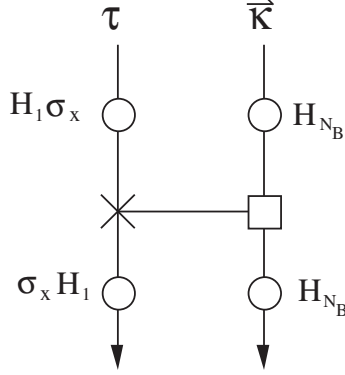


Figure 2: Qubit circuit for D-J's algorithm.

where $\vec{n}(\vec{\kappa}) = (n(\kappa_0), n(\kappa_1), \dots, n(\kappa_{N_B-1}))$. The operation $\sigma_x^{f(\vec{n}(\vec{\kappa}))}(\tau)$, because it depends on f , is often called an “oracle” and each use of it is called a “query”. The right hand side of Eq.(25) may be represented by the circuit diagram shown in Fig.2. The D-J algorithm consists of applying Ω to an initial state $|0\rangle_{\vec{\kappa}}|0\rangle_{\tau}$ of bits $\vec{\kappa}$ and τ , and then measuring the final state of these bits in the preferred basis.

Fig.2 and the right hand side of Eq.(25) are two equivalent ways of representing a particular SEO. There are infinitely many SEOs that yield Ω . Fig.2 is just one of them. In fact, the original D-J paper[2] gave a different SEO for Ω , one with two queries instead of one.

For $X \in Bool^{N_B}, Y \in Bool$, let

$$|\psi_0\rangle = |X\rangle_{\vec{\kappa}}|Y\rangle_{\tau}, \quad (26)$$

and

$$|\psi_i\rangle = \Omega_i|\psi_{i-1}\rangle \text{ for } i = 1, 2, \dots, \quad (27)$$

where

$$\Omega_1 = \hat{H}_1(\vec{\kappa})\hat{H}_1(\tau)\sigma_x(\tau), \quad (28)$$

$$\Omega_2 = \sigma_x^{f(\vec{n}(\vec{\kappa}))}(\tau), \quad (29)$$

and

$$\Omega_3 = \Omega_1^\dagger. \quad (30)$$

Then it is easy to show using simple identities (such as $(H_1)_{b,b'} = (-1)^{bb'}$, $\bar{0} = 1$, $\bar{1} = 0$, and $(-1)^b = (-1)^{-b}$ for $b, b' \in Bool$) that

$$|\psi_1\rangle = \frac{1}{\sqrt{2^{N_B+1}}} \sum_{x,y} (-1)^{x \cdot X + y \bar{Y}} |x\rangle_{\vec{\kappa}} |y\rangle_{\tau} , \quad (31)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2^{N_B+1}}} \sum_{x,y} (-1)^{x \cdot X + y \bar{Y}} |x\rangle_{\vec{\kappa}} |y \oplus f(x)\rangle_{\tau} , \quad (32)$$

$$|\psi_3\rangle = \frac{1}{2^{N_B+1}} \sum_{x,y,X',Y'} (-1)^{x \cdot (X' - X) + y(\bar{Y}' - \bar{Y}) + \bar{Y}' f(x)} |X'\rangle_{\vec{\kappa}} |Y'\rangle_{\tau} . \quad (33)$$

Applying $\langle X', Y' |$ to the right hand side of Eq.(33) and using the identity Eq.(1) finally yields:

$$\langle X', Y' | \Omega | X, Y \rangle = \delta(Y', Y) \frac{1}{2^{N_B}} \sum_{x \in Bool^{N_B}} (-1)^{x \cdot (X' - X) + \bar{Y}' f(x)} \quad (34)$$

for all $X', X \in Bool^{N_B}$ and $Y', Y \in Bool$. Thus, if the initial states of $\vec{\kappa}$ and τ are $\underline{X} = 0$ and $\underline{Y} = 0$, then the probability of obtaining $\underline{X}' = X'$ for the final state of $\vec{\kappa}$ is

$$\begin{aligned} P(X' | \underline{X} = \underline{Y} = 0) &= \sum_{Y'} |\langle X', Y' | \Omega | X = 0, Y = 0 \rangle|^2 \\ &= \frac{1}{4^{N_B}} \left| \sum_x (-1)^{x \cdot X' + f(x)} \right|^2 . \end{aligned} \quad (35)$$

Let \mathcal{F}_{bal} , the set of “balanced” functions, be the set of all $f : Bool^{N_B} \rightarrow Bool$ such that f maps exactly half of its domain to zero and half to one. Let \mathcal{F}_{con} , the set of “constant” functions, be the set of all $f : Bool^{N_B} \rightarrow Bool$ such that f maps all its domain to zero or all of it to one. From Eq.(35), if $X' = 0$ and $f \in \mathcal{F}_{bal} \cup \mathcal{F}_{con}$, then

$$P(\underline{X}' = 0 | \underline{X} = \underline{Y} = 0) = \begin{cases} 1 & \text{if } f \in \mathcal{F}_{con} \\ 0 & \text{if } f \in \mathcal{F}_{bal} \end{cases} . \quad (36)$$

Now consider the QB net defined by Fig.3 and Table 1.

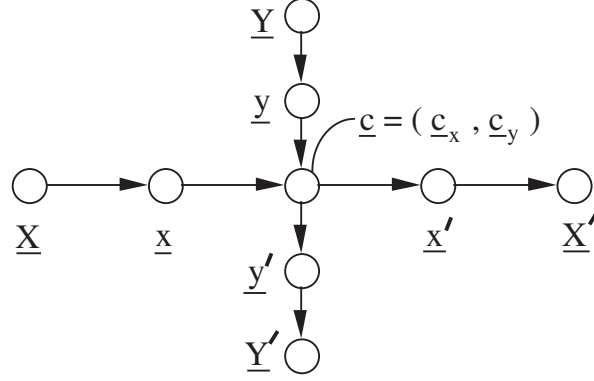


Figure 3: QB net for D-J's algorithm.

nodes	states	amplitudes	comments
\underline{X}	$X \in Bool^{N_B}$	$\delta(X, 0)$	
\underline{Y}	$Y \in Bool$	$\delta(Y, 0)$	
\underline{x}	$x \in Bool^{N_B}$	$(-1)^{x \cdot X} / \sqrt{2^{N_B}}$	
\underline{y}	$y \in Bool$	$(-1)^{y \bar{Y}} / \sqrt{2}$	
\underline{c}	$c = (c_x, c_y), c_x \in Bool^{N_B}, c_y \in Bool$	$\delta(c_x, x) \delta(c_y, y \oplus f(x))$	
$\underline{x'}$	$x' \in Bool^{N_B}$	$\delta(x', c_x)$	
$\underline{y'}$	$y' \in Bool$	$\delta(y', c_y)$	
$\underline{X'}$	$X' \in Bool^{N_B}$	$(-1)^{X' \cdot x'} / \sqrt{2^{N_B}}$	
$\underline{Y'}$	$Y' \in Bool$	$(-1)^{\bar{Y'} y'} / \sqrt{2}$	

Table 1

For this net, the amplitude $A(x.)$ of net story $x.$ is the product of all the terms in the third column of Table 1. If $\underline{X} = 0$ and $\underline{Y} = 0$, then the probability of obtaining $\underline{X'} = X'$ is

$$P(X' | \underline{X} = \underline{Y} = 0) = \frac{\sum_{Y'} \left| \sum_{all/X'Y', X, Y} A(x.) \right|_{X=Y=0}^2}{\sum_{X'} num}, \quad (37)$$

where $A(x.)$ on the right hand side is evaluated at $X = Y = 0$. Substituting the value of $A(x.)$ into Eq.(37) immediately yields Eq.(35).

Note that one can calculate the probability distribution Eq.(35) by means of a CB net instead of a QB net. One can do this with the CB net defined by the graph $\underline{X'} \rightarrow \underline{Y'}$, with:

nodes	states	probabilities	comments
\underline{X}'	$X' \in Bool^{N_B}$	$P_{\underline{X}'}(X')$	
\underline{Y}'	$Y' \in Bool$	$P_{\underline{Y}' \underline{X}'}(Y' X')$	

Table 2

where $P_{\underline{X}'}$ and $P_{\underline{Y}'|\underline{X}'}$ are calculated from

$$P_{\underline{X}', \underline{Y}'}(X', Y') = |\langle X', Y' | \Omega | X = 0, Y = 0 \rangle|^2. \quad (38)$$

We will say that the CB net defined by the graph $\underline{X}' \rightarrow \underline{Y}'$ and Table 2 is “q-embedded” in the QB net defined by Fig.3 and Table 1. In subsequent sections, we will say much more about q-embedding of CB nets.

3.2 Simon’s Algorithm

In this section we will discuss Simon’s algorithm[3]. We will do this first in terms of qubit circuits (the conventional approach), and then in terms of QB nets.

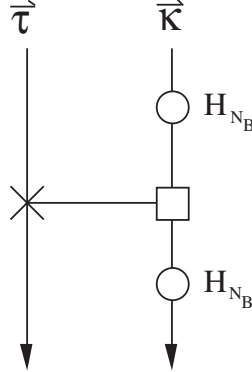


Figure 4: Qubit circuit for Simon’s algorithm.

Simon’s algorithm uses N_B “control” bits, just like the D-J algorithm. However, it uses N_B target bits whereas the D-J algorithm uses only one. Simon’s algorithm deals with a vector-valued function $f : Bool^{N_B} \rightarrow Bool^{N_B}$, whereas D-J’s algorithm deals with a scalar-valued function $f : Bool^{N_B} \rightarrow Bool$. Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B “control” bits and let $\vec{\tau} = (\tau_0, \tau_1, \dots, \tau_{N_B-1})$ label N_B “target” bits. Assume all τ_i and κ_i are distinct. We will denote the state of these bits in the preferred basis (the eigenvectors of σ_z) by $|x\rangle_{\vec{\kappa}}|y\rangle_{\vec{\tau}}$, where $x \in Bool^{N_B}$ and $y \in Bool^{N_B}$. Given a function $f = (f_0, f_1, \dots, f_{N_B-1})$ where $f_i : Bool^{N_B} \rightarrow Bool$, define the unitary operator Ω by

$$\Omega = \hat{H}_1(\vec{\kappa}) \left[\prod_{i=0}^{N_B-1} \sigma_x^{f_i(\vec{n}(\vec{\kappa}))}(\tau_i) \right] \hat{H}_1(\vec{\kappa}) . \quad (39)$$

The operator Ω for Simon's algorithm is analogous to the Ω defined by Eq.(25) for the D-J algorithm. The right hand side of Eq.(39) may be represented by the circuit diagram of Fig.4. Simon's algorithm consists of applying Ω given by Eq.(39) to an initial state $|0\rangle_{\vec{\kappa}}|0\rangle_{\vec{\tau}}$ of bits $\vec{\kappa}$ and $\vec{\tau}$, and then measuring the final state of these bits in the preferred basis. One performs this routine several times. The measurement outcomes allow one to determine the period of the function f if f is of a special periodic type that will be specified later.

Using the same techniques that we used to evaluate the matrix elements of Ω for the D-J algorithm, one finds

$$\langle X', Y' | \Omega | X, Y \rangle = \frac{1}{2^{N_B}} \sum_{x \in \text{Bool}^{N_B}} (-1)^{x \cdot (X' - X)} \delta(Y', Y \oplus f(x)) , \quad (40)$$

for all $X', Y', X, Y \in \text{Bool}^{N_B}$. If the initial states of $\vec{\kappa}$ and $\vec{\tau}$ are $\underline{X} = 0$ and $\underline{Y} = 0$, then the probability of obtaining $\underline{X}' = X'$ for the final state of $\vec{\kappa}$ is

$$\begin{aligned} P(X' | \underline{X} = \underline{Y} = 0) &= \sum_{Y'} |\langle X', Y' | \Omega | X = 0, Y = 0 \rangle|^2 \\ &= \frac{1}{4^{N_B}} \sum_{Y'} \left| \sum_x (-1)^{x \cdot X'} \delta(Y', f(x)) \right|^2 . \end{aligned} \quad (41)$$

Now suppose \mathcal{F}_S is the set of those functions $f : \text{Bool}^{N_B} \rightarrow \text{Bool}^{N_B}$ such that f is 2 to 1 (i.e., f maps exactly two domain points into each image point) and has a "period" Δ . By a period Δ , we mean a non-zero element of Bool^{N_B} such that $f(x) = f(x \oplus \Delta)$ for all $x \in \text{Bool}^{N_B}$. For any $f \in \mathcal{F}_S$ and any $y \in \text{Bool}^{N_B}$, there exist exactly two elements of Bool^{N_B} , call them x_1 and x_2 , such that $x_1 = x_2 \oplus \Delta$ and $f(x_1) = f(x_2) = y$. Call $f_p^{-1}(y)$ one of these x values, and call $f_p^{-1}(y) \oplus \Delta$ the other. (The p subscript stands for "principal part", in analogy with Complex Analysis.) If $f \in \mathcal{F}_S$, and $I(f)$ is the image of f , then

$$\delta(Y', f(x)) = \begin{cases} \delta(f_p^{-1}(Y'), x) + \delta(f_p^{-1}(Y') \oplus \Delta, x), & \text{if } Y' \in I(f) \\ 0 & \text{otherwise} \end{cases} . \quad (42)$$

Substituting this expression for $\delta(Y', f(x))$ into Eq.(41) and using Eq.(1) yields

$$P(X' | \underline{X} = \underline{Y} = 0) = \frac{1}{2^{N_B-1}} \delta(X' \cdot \Delta, 0) . \quad (43)$$

To calculate the period Δ of f , run the experiment ν times, measuring X' each time. Let $X'(i)$ represent the i th measurement outcome. Then, for sufficiently large ν , one can find Δ by solving the equations $X'(1) \cdot \Delta = 0$, $X'(2) \cdot \Delta = 0$, ..., $X'(\nu) \cdot \Delta = 0$.

Now consider the QB net defined by Fig.5 and Table 3.

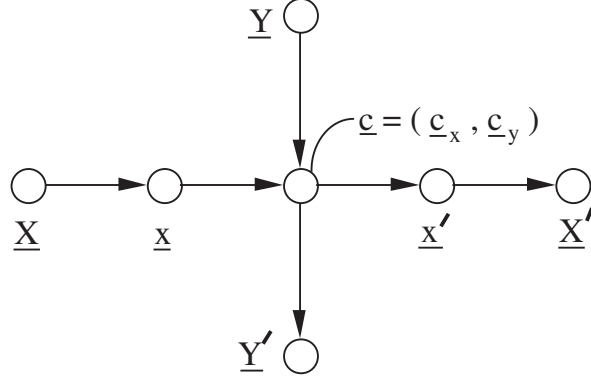


Figure 5: QB net for Simon's algorithm.

nodes	states	amplitudes	comments
\underline{X}	$X \in Bool^{N_B}$	$\delta(X, 0)$	
\underline{Y}	$Y \in Bool^{N_B}$	$\delta(Y, 0)$	
\underline{x}	$x \in Bool^{N_B}$	$(-1)^{x \cdot X} / \sqrt{2^{N_B}}$	
\underline{c}	$c = (c_x, c_y); c_x, c_y \in Bool^{N_B}$	$\delta(c_x, x) \delta(c_y, Y \oplus f(x))$	
$\underline{x'}$	$x' \in Bool^{N_B}$	$\delta(x', c_x)$	
$\underline{X'}$	$X' \in Bool^{N_B}$	$(-1)^{X' \cdot x'} / \sqrt{2^{N_B}}$	
$\underline{Y'}$	$Y' \in Bool^{N_B}$	$\delta(Y', c_y)$	

Table 3

For this net, the amplitude $A(x.)$ of net story $x.$ is the product of all the terms in the third column of Table 3. If $\underline{X} = 0$ and $\underline{Y} = 0$, then the probability of obtaining $\underline{X'} = X'$ is

$$P(X' | \underline{X} = \underline{Y} = 0) = \frac{\sum_{Y'} \left| \sum_{all/X'Y', X, Y} A(x.) \right|_{X=Y=0}^2}{\sum_{X'} num}, \quad (44)$$

where $A(x.)$ on the right hand side is evaluated at $X = Y = 0$. Substituting the value of $A(x.)$ into Eq.(44) immediately yields Eq.(41).

It is possible to calculate the probability distribution Eq.(41) by means of a CB net instead of a QB net. One can do this with the CB net defined by the graph $\underline{X'} \rightarrow \underline{Y'}$, with:

nodes	states	probabilities	comments
\underline{X}'	$X' \in Bool^{N_B}$	$P_{\underline{X}'}(X')$	
\underline{Y}'	$Y' \in Bool^{N_B}$	$P_{\underline{Y}' \underline{X}'}(Y' X')$	

Table 4

where $P_{\underline{X}'}$ and $P_{\underline{Y}'|\underline{X}'}$ are calculated from

$$P_{\underline{X}',\underline{Y}'}(X', Y') = |\langle X', Y' | \Omega | X = 0, Y = 0 \rangle|^2 . \quad (45)$$

We will say that the CB net defined by the graph $\underline{X}' \rightarrow \underline{Y}'$ and Table 4 is “q-embedded” in the QB net defined by Fig.5 and Table 3.

3.3 Bernstein-Vazirani Algorithm

In this section we will discuss the B-V (Bernstein-Vazirani) algorithm[4].

To understand the B-V algorithm, it is helpful to first establish the following simple single qubit identities. First note that the single qubit Hadamard matrix rotates the Z -direction number operator into the X -direction number operator:

$$\hat{H}_1 n_z \hat{H}_1 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = n_x . \quad (46)$$

Thus,

$$\sigma_x^b = [(-1)^{n_x}]^b = (-1)^{bn_x} = \hat{H}_1 (-1)^{bn_z} \hat{H}_1 . \quad (47)$$

Next note that σ_x exchanges the components of any vector it acts on:

$$\sigma_x \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} , \quad (48)$$

for any complex numbers α, β . In particular, if $b \in Bool$, then

$$\sigma_x^b |0\rangle = |b\rangle . \quad (49)$$

Now we are ready to discuss the B-V algorithm. Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B “control” bits and let τ label a single “target” bit. Assume that τ and all the κ_i are distinct. We will denote the state of these bits in the preferred basis (the eigenvectors of σ_z) by $|x\rangle_{\vec{\kappa}} |y\rangle_{\tau}$, where $x \in Bool^{N_B}$ and $y \in Bool$. For $\vec{b} \in Bool^{N_B}$, define the unitary operator

$$\omega_{\vec{b}} = \prod_{i=0}^{N_B-1} \sigma_x(\kappa_i)^{b_i} . \quad (50)$$

The B-V algorithm is simply the following multi-qubit generalization of Eq.(49)

$$\omega_{\vec{b}}|0\rangle_{\vec{\kappa}} = |\vec{b}\rangle_{\vec{\kappa}} . \quad (51)$$

That's all there is to B-V!

Eq.(51) can be represented by a qubit circuit consisting of a single wire for $\vec{\kappa}$, with a single node representing $\omega_{\vec{b}}$. Eq.(51) can also be represented by a QB net defined by the graph $\underline{X} \rightarrow \underline{X}'$, with

nodes	states	amplitudes	comments
\underline{X}	$X = (X_0, X_1, X_{N_B-1}) \in Bool^{N_B}$	$\delta(X, 0)$	
\underline{X}'	$X' = (X'_0, X'_1, X'_{N_B-1}) \in Bool^{N_B}$	$\prod_{i=0}^{N_B-1} \delta^{b_i}(X'_i, \overline{X_i})$	

Table 5

We should mention that it is common in the literature to dress up and obfuscate Eq.(50) as follows. By virtue of Eq.(47), one can re-express $\omega_{\vec{b}}$ as

$$\omega_{\vec{b}} = \prod_{i=0}^{N_B-1} \sigma_x(\kappa_i)^{b_i} = \hat{H}_1(\vec{\kappa})(-1)^{\sum_{i=0}^{N_B-1} b_i n_z(\kappa_i)} \hat{H}_1(\vec{\kappa}) . \quad (52)$$

Some workers ascend to an even higher peak of obfuscation by adding a totally unnecessary target qubit. They define an operator, call it $\Omega_{\vec{b}}$, obtained by replacing the (-1) in Eq.(52) by the operator $\sigma_x(\tau)$ acting on a target qubit τ :

$$\Omega_{\vec{b}} = \hat{H}_1(\vec{\kappa})[\sigma_x(\tau)]^{\sum_{i=0}^{N_B-1} b_i n_z(\kappa_i)} \hat{H}_1(\vec{\kappa}) . \quad (53)$$

At the beginning of the experiment, they put the target qubit in a state which is an eigenvector of $\sigma_x(\tau)$ with eigenvalue -1 . Thus, the obfuscated version of the B-V algorithm with a target qubit can be summarized by

$$\Omega_{\vec{b}}|{-x}\rangle_{\tau}|0\rangle_{\vec{\kappa}} = \omega_{\vec{b}}|{-x}\rangle_{\tau}|0\rangle_{\vec{\kappa}} = |{-x}\rangle_{\tau}|\vec{b}\rangle_{\vec{\kappa}} . \quad (54)$$

We emphasize that for the B-V algorithm, the target qubit is a totally unnecessary affectation.

So far we have given an unconventional presentation of the B-V algorithm. For completeness, we now give a conventional one. Define

$$|\psi_0\rangle = |0\rangle_{\vec{\kappa}}|{-x}\rangle_{\tau} , \quad (55)$$

and

$$|\psi_i\rangle = \Omega_i|\psi_{i-1}\rangle \text{ for } i = 1, 2, \dots , \quad (56)$$

where

$$\Omega_1 = \hat{H}_1(\vec{\kappa}) , \quad (57)$$

$$\Omega_2 = [\sigma_x(\tau)]^{\sum_{i=0}^{N_B-1} b_i n_z(\kappa_i)}, \quad (58)$$

and

$$\Omega_3 = \Omega_1. \quad (59)$$

It follows that

$$|\psi_1\rangle = \frac{1}{\sqrt{2^{N_B}}} \sum_{\vec{x} \in \text{Bool}^{N_B}} |\vec{x}\rangle_{\vec{\kappa}} |-x\rangle_{\tau}, \quad (60)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2^{N_B}}} \sum_{\vec{x}} (-1)^{\vec{b} \cdot \vec{x}} |\vec{x}\rangle_{\vec{\kappa}} |-x\rangle_{\tau}, \quad (61)$$

and

$$|\psi_3\rangle = \frac{1}{\sqrt{2^{N_B}}} \sum_{\vec{x}} (-1)^{\vec{b} \cdot \vec{x}} \frac{1}{\sqrt{2^{N_B}}} \sum_{\vec{y}} (-1)^{\vec{y} \cdot \vec{x}} |\vec{y}\rangle_{\vec{\kappa}} |-x\rangle_{\tau} \quad (62a)$$

$$= \frac{1}{2^{N_B}} \sum_{\vec{x}, \vec{y}} (-1)^{(\vec{b} - \vec{y}) \cdot \vec{x}} |\vec{y}\rangle_{\vec{\kappa}} |-x\rangle_{\tau} \quad (62b)$$

$$= |\vec{b}\rangle_{\vec{\kappa}} |-x\rangle_{\tau}. \quad (62c)$$

To go from step (b) to step (c) of Eq.(62), we used the orthogonality property given by Eq.(1).

3.4 Grover's Algorithm

In this section we will discuss Grover's algorithm [5].

Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B bits. Assume all κ_i are distinct. We begin by defining the following N_S -dimensional column vectors:

$$|\mu\rangle_{\vec{\kappa}} = \mu = \mu_{N_S} = \frac{1}{\sqrt{N_S}} [1, 1, 1, 1, \dots, 1]^T = \frac{1}{\sqrt{N_S}} H_{N_B} [1, 0, 0, 0, \dots, 0]^T, \quad (63)$$

$$|\phi\rangle_{\vec{\kappa}} = \phi = [0, \dots, 0, 0, 1, 0, 0, \dots, 0]^T. \quad (64)$$

All components of ϕ are zero except for one predetermined component, located at position $j_{\text{targ}} \in Z_{0, N_S-1}$, which equals one. We will refer to j_{targ} as the **target state** (not to be confused with a target qubit). Note that we chose a special basis (or, equivalently, a special matrix representation) from the start. Note that $\langle \phi | \mu \rangle = \frac{1}{\sqrt{N_S}}$, so μ and ϕ are nearly orthogonal for large N_S . It is also convenient to define the component-wise negation of ϕ :

$$|\phi^{not}\rangle_{\vec{\kappa}} = \phi^{not} = [1, \dots, 1, 1, 0, 1, 1, \dots, 1]^T . \quad (65)$$

Note that ϕ^{not} is not normalized.

Define projection and reflection operators for μ and ϕ :

$$\Pi_{\mu} = |\mu\rangle\langle\mu| , \quad R_{\mu} = 1 - 2\Pi_{\mu} = (-1)^{\Pi_{\mu}} , \quad (66)$$

and

$$\Pi_{\phi} = |\phi\rangle\langle\phi| , \quad R_{\phi} = 1 - 2\Pi_{\phi} = (-1)^{\Pi_{\phi}} . \quad (67)$$

Grover's algorithm can be summarized by the following equation[10][11]:

$$(-R_{\mu}R_{\phi})^r \mu \approx \phi , \quad (68)$$

for some integer r to be determined, where “ \approx ” means approximation at large N_S . Thus, starting with an N_B qubit system in a state μ , one applies the operator $(-R_{\mu}R_{\phi})$ consecutively r times, so that the N_B qubit system ends in a state as close to ϕ as possible. Measuring state ϕ in the special basis yields the target state j_{targ} .

Eq.(68) can be represented by a qubit circuit consisting of a single wire for $\vec{\kappa}$, with r nodes, each representing $-R_{\mu}R_{\phi}$. Eq.(68) can also be represented by a QB net defined by a Markov chain graph $\underline{X}_0 \rightarrow \underline{X}_1 \rightarrow \underline{X}_2 \rightarrow \dots \rightarrow \underline{X}_{r-1}$, with

nodes	states	amplitudes	comments
\underline{X}_0	$X_0 \in Bool^{N_B}$	$\delta(X_0, 0)$	
\underline{X}_i for $i \in Z_{1,r-1}$	$X_i \in Bool^{N_B}$	$\langle X_i (-R_{\mu}R_{\phi}) X_{i-1} \rangle$	

Table 6

To find the optimum number r of iterations, one can proceed as follows.

First, notice that Eq.(68) describes a process which is entirely confined to the vector subspace spanned by μ and ϕ . Since μ and ϕ are not orthogonal, it is convenient to define an orthonormal basis e_0, e_1 for the space $span(\mu, \phi)$. Let

$$e_0 = \phi , \quad e_1 = \frac{\phi^{not}}{\sqrt{N_S - 1}} . \quad (69)$$

Then

$$\mu = \frac{1}{\sqrt{N_S}}(e_0 + \sqrt{N_S - 1} e_1) . \quad (70)$$

Fig.6 portrays various vectors that arise in explaining Grover's algorithm.

Since we plan to stay within the two dimensional vector space with orthonormal basis e_0, e_1 , it is convenient to switch matrix representations. Within $span(e_0, e_1)$, e_0, e_1 can be represented more simply by:

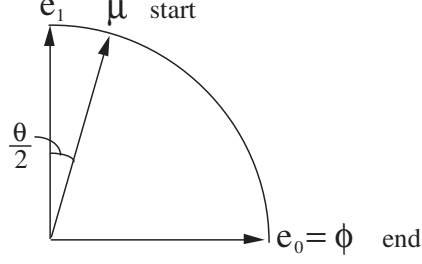


Figure 6: Various vectors relevant to Grover's Algorithm.

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (71)$$

If e_0, e_1 are represented in this way, then

$$\phi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mu = \frac{1}{\sqrt{N_S}} \begin{pmatrix} 1 \\ \sqrt{N_S - 1} \end{pmatrix}, \quad (72)$$

and

$$-R_\mu R_\phi = \frac{1}{N_S} \begin{pmatrix} N_S - 2 & 2\sqrt{N_S - 1} \\ -2\sqrt{N_S - 1} & N_S - 2 \end{pmatrix}. \quad (73)$$

Thus,

$$-R_\mu R_\phi = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (74)$$

where

$$\sin \theta = \frac{2\sqrt{N_S - 1}}{N_S} \approx \frac{2}{\sqrt{N_S}}. \quad (75)$$

Eq.(74) is just Eq.(24) with $e'_1 = \mu$ and $e_0 = \phi$. It follows that

$$(-R_\mu R_\phi)^r = \begin{pmatrix} \cos(r\theta) & \sin(r\theta) \\ -\sin(r\theta) & \cos(r\theta) \end{pmatrix}, \quad (76)$$

and

$$\begin{aligned} (-R_\mu R_\phi)^r \mu &= \begin{pmatrix} \cos(r\theta) & \sin(r\theta) \\ -\sin(r\theta) & \cos(r\theta) \end{pmatrix} \frac{1}{\sqrt{N_S}} \begin{pmatrix} 1 \\ \sqrt{N_S - 1} \end{pmatrix} \\ &\approx \begin{pmatrix} \sin(r\theta) \\ \cos(r\theta) \end{pmatrix}. \end{aligned} \quad (77)$$

We want the final state of the system to be parallel or anti-parallel to $e_0 = \phi$; therefore, we want

$$\begin{pmatrix} \sin(r\theta) \\ \cos(r\theta) \end{pmatrix} \approx \begin{pmatrix} \pm 1 \\ 0 \end{pmatrix} . \quad (78)$$

This will occur if

$$r\theta \approx \frac{\pi}{2}(1 + 2k) \quad , \quad r \approx \frac{\pi}{4}(1 + 2k)\sqrt{N_S} \quad (79)$$

for some integer k .

Note that, in Grover's algorithm, the number of "queries" (calls to a unitary matrix that depends on ϕ) is far from unique. To illustrate this, let Q be a permutation matrix that satisfies

$$Q\phi = |0\rangle = [1, 0, 0, \dots, 0]^T . \quad (80)$$

Since all the components of μ are equal, $Q\mu = \mu$. Thus

$$(-R_\mu R_\phi)^r \mu = Q^T (-R_\mu R_{|0\rangle})^r Q\mu = Q^T (-R_\mu R_{|0\rangle})^r \mu . \quad (81)$$

Hence, it is possible to accomplish the full Grover transformation of μ with only a single query Q^T .

Since $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix}$, the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is just a clockwise rotation by $\pi/2$. Let

$$\begin{aligned} U_{Grover} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= -e_1 e_0^T + e_0 e_1^T \\ &= \frac{1}{\sqrt{N_S - 1}} [-\phi^{not} \phi^T + \phi(\phi^{not})^T] . \end{aligned} \quad (82)$$

Note that

$$\begin{aligned} U_{Grover} \mu &= \frac{1}{\sqrt{N_S - 1}} (-\phi^{not} [\phi^T \mu] + \phi [(\phi^{not})^T \mu]) \\ &= \frac{1}{\sqrt{N_S(N_S - 1)}} [-\phi^{not} + (N_S - 1)\phi] . \end{aligned} \quad (83)$$

From the point of view of quantum compiling, what Grover found is that the $\pi/2$ rotation U_{Grover} is (approximately) equal to the r -fold product of $-R_\mu R_\phi$, where $-R_\mu R_\phi$ can be shown to have a SEO of low (polynomial in N_B) complexity.

Grover's algorithm has been modified in various, minor ways since it was first published. For example, Brassard et al. pointed out in Ref.[12] that the vector μ need not be the vector whose components are all equal. Other vectors μ will do just as well. Another modification of Grover's algorithm due to Younes-Miller[13] adds an extra qubit to the original N_B qubits. Next we will discuss the Younes-Miller modification of Grover's algorithm, because it resembles a modification of Grover's algorithm that we will use in a future section.

Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B bits. Let τ label a single bit. Assume τ and all the κ_i are distinct. Let μ and ϕ denote the same N_S dimensional column vectors that we used in discussing the original Grover algorithm. In addition, define the following $2N_S$ dimensional column vectors:

$$|\tilde{\mu}\rangle = |+_z\rangle_\tau |\mu\rangle_{\vec{\kappa}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \mu_{N_S} = \begin{pmatrix} \mu_{N_S} \\ 0 \end{pmatrix}, \quad (84)$$

$$|\tilde{\phi}\rangle = |-_x\rangle_\tau |\phi\rangle_{\vec{\kappa}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi \\ -\phi \end{pmatrix}. \quad (85)$$

Note that $\langle \tilde{\phi} | \tilde{\mu} \rangle = \frac{1}{\sqrt{2N_S}}$, so $\tilde{\phi}$ and $\tilde{\mu}$ are nearly orthogonal for large N_S . Define projection and reflection operators for $\tilde{\phi}$ in the usual way:

$$\Pi_{\tilde{\phi}} = |\tilde{\phi}\rangle\langle\tilde{\phi}|, \quad R_{\tilde{\phi}} = 1 - 2\Pi_{\tilde{\phi}} = 1 - 2\Pi_{|\phi\rangle_{\vec{\kappa}}} \Pi_{|-_x\rangle_\tau}. \quad (86)$$

$R_{\tilde{\phi}}$ can be re-expressed as

$$\begin{aligned} R_{\tilde{\phi}} &= 1 + \Pi_{|\phi\rangle_{\vec{\kappa}}}(\sigma_x(\tau) - 1) = \exp[\Pi_{|\phi\rangle_{\vec{\kappa}}} \ln \sigma_x(\tau)] = \\ &= [\sigma_x(\tau)]^{\Pi_{|\phi\rangle_{\vec{\kappa}}}}. \end{aligned} \quad (87)$$

Define projection and reflection operators for $\tilde{\mu}$ in the usual way:

$$\Pi_{\tilde{\mu}} = |\tilde{\mu}\rangle\langle\tilde{\mu}|, \quad R_{\tilde{\mu}} = 1 - 2\Pi_{\tilde{\mu}} = 1 - 2\Pi_{|\mu\rangle_{\vec{\kappa}}} \Pi_{|+_z\rangle_\tau}. \quad (88)$$

$R_{\tilde{\mu}}$ can be re-expressed as

$$R_{\tilde{\mu}} = \hat{H}_1(\vec{\kappa}) \left(1 - 2\Pi_{|0\rangle_{\vec{\kappa}}} \Pi_{|0\rangle_\tau}\right) \hat{H}_1(\vec{\kappa}). \quad (89)$$

In analogy with the original Grover's algorithm, the Younes-Miller version can be summarized by

$$(-R_{\tilde{\mu}} R_{\tilde{\phi}})^r |\tilde{\mu}\rangle \approx |\tilde{\phi}\rangle, \quad (90)$$

for some integer r to be determined, where “ \approx ” means approximation at large N_S . Thus, starting with an $N_B + 1$ qubit system in a state $\tilde{\mu}$, one applies the operator $(-R_{\tilde{\mu}} R_{\tilde{\phi}})$ consecutively r times, so that the final state of the $N_B + 1$ qubit system

ends in a state as close to $\tilde{\phi}$ as possible. Measuring state $\tilde{\phi}$ in the special basis yields the target state j_{targ} .

To find the optimum number r of iterations, one can proceed as follows.

First, notice that Eq.(90) describes a process which is entirely confined to the vector subspace spanned by $\tilde{\mu}$ and $\tilde{\phi}$. Since $\tilde{\mu}$ and $\tilde{\phi}$ are not orthogonal, it is convenient to define an orthonormal basis e_0, e_1 for the space $span(\tilde{\mu}, \tilde{\phi})$. Let

$$e_0 = \tilde{\phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi \\ -\phi \end{pmatrix}, \quad (91)$$

and

$$e_1 = \frac{1}{K} [\tilde{\mu} - (\tilde{\mu} \cdot e_0) e_0], \quad (92)$$

where K is chosen so that $e_1^2 = 1$. It is easy to show that

$$K = |\tilde{\mu} - (\tilde{\mu} \cdot e_0) e_0| = \sqrt{\frac{N_S - \frac{1}{2}}{N_S}}. \quad (93)$$

Thus,

$$e_1 = \frac{1}{\sqrt{N_S - \frac{1}{2}}} \begin{pmatrix} \phi^{not} + \frac{\phi}{2} \\ \frac{\phi}{2} \end{pmatrix}. \quad (94)$$

Furthermore,

$$\tilde{\mu} = \frac{1}{\sqrt{2N_S}} [e_0 + \sqrt{2N_S - 1} e_1]. \quad (95)$$

Fig.7 portrays various vectors that arise in explaining Younes' version of Grover's algorithm.

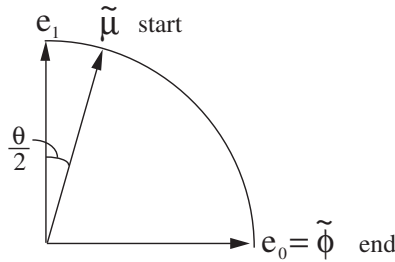


Figure 7: Various vectors relevant to Younes' version of Grover's Algorithm.

Since we plan to stay within the two dimensional vector space with orthonormal basis e_0, e_1 , it is convenient to switch matrix representations. Within $span(e_0, e_1)$, e_0, e_1 can be represented more simply by:

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (96)$$

If e_0, e_1 are represented this way, then

$$\tilde{\phi} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \tilde{\mu} = \frac{1}{\sqrt{2N_S}} \begin{pmatrix} 1 \\ \sqrt{2N_S - 1} \end{pmatrix}, \quad (97)$$

and

$$-R_{\tilde{\mu}}R_{\tilde{\phi}} = \frac{1}{N_S} \begin{pmatrix} N_S - 1 & \sqrt{2N_S - 1} \\ -\sqrt{2N_S - 1} & N_S - 1 \end{pmatrix}. \quad (98)$$

Thus,

$$-R_{\tilde{\mu}}R_{\tilde{\phi}} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (99)$$

where

$$\sin \theta = \frac{\sqrt{2N_S - 1}}{N_S} \approx \sqrt{\frac{2}{N_S}}. \quad (100)$$

A comparison of Eq.(72) (for the original Grover's algorithm) and Eq.(97) (for Younes's version of Grover's algorithm) reveals that for the purpose of finding the optimal number r of iterations, Younes' algorithm is the same as Grover's algorithm if one replaces N_S in Grover's algorithm by $2N_S$. This comes from the fact that Younes' algorithm uses $N_B + 1$ bits whereas Grover's uses N_B .

4 Generalization of Standard Algorithms, a list of Desiderata

So far we have analyzed several standard quantum computing algorithms, namely those attributed to Deutsch-Jozsa, Bernstein-Vazirani, Simon and Grover. (Two other standard algorithms that we didn't analyze are Shor's algorithm[14] and the algorithm for Teleportation[15].) In this section, we will try to point out those features of the standard algorithms that would be, in our opinion, fruitful to generalize. Bear in mind that generalizations are seldom unique, but some are more natural, fruitful and far-reaching than others.

(a) Allow more complicated graph topologies

The standard algorithms discussed here can all be represented by QB nets with trivial topologies such as 2 body scattering graphs or Markov chains. However,

other important quantum algorithms, such as the one for Teleportation[15], can be represented by QB nets with more complicated graph topologies (e.g., with loops).

(b) Estimate more general probability distributions

The goal of most standard algorithms is to estimate a deterministic probability distribution. However, estimating non-deterministic ones is also very useful. Such estimates are useful in, for example, applications of Decision Theory and Artificial Intelligence, where inferences are made based on uncertain knowledge.

(c) Allow multiple runs and the rejection of some

If one is estimating a non-deterministic probability distribution, it will be necessary to do multiple runs. It may also be necessary to allow rejection of runs. Obviously, the number of rejected runs is best kept as small as possible.

(d) Allow more general measurements

Suppose \underline{x} is a node of a QB net. Let $S_{\underline{x}}$ be the set of its possible states. We will say that node \underline{x} has been measured if during the experiment which the QB net describes, a measurement is performed that restricts the possible states of \underline{x} to a proper subset $S'_{\underline{x}}$ of $S_{\underline{x}}$. When \underline{x} is an internal (ditto, external) node of the QB net, we will refer to its measurement as an internal (ditto, external) measurement.

The standard algorithms discussed here use external but no internal measurements. However, other important quantum algorithms, such as the one for Teleportation, do use internal ones.

5 Q-Embeddings

The remainder of this paper will be devoted to discussing a class of algorithms which generalizes some standard algorithms and achieves some of the desiderata given in the previous section. Our algorithms are based on the idea that, given a CB net, one can always embed it in a QB net. Simple examples of such q-embeddings have already been given in the sections dealing with standard algorithms.

We start by defining some terminology that will be useful.

A **probability matrix** $P(y|x)$ is a rectangular (not necessarily square) matrix with row index $y \in S_{\underline{y}}$ and column index $x \in S_{\underline{x}}$ such that $P(y|x) \geq 0$ for all x, y , and $\sum_y P(y|x) = 1$ for all x . The set of all probability matrices $P(y|x)$ where $x \in S_{\underline{x}}$ and $y \in S_{\underline{y}}$ will be denoted by $pd(S_{\underline{y}}|S_{\underline{x}})$ (pd = probability distribution). A probability matrix is assigned to each node of a CB net. A **probability matrix** $P(y|x)$ is **deterministic** if for each column x , there exists a single row y , call it

$y(x)$, such that $P(y|x) = \delta(y(x), y)$. Any map $f : S_{\underline{x}} \rightarrow S_{\underline{y}}$ uniquely specifies (and is uniquely specified) by the deterministic probability matrix P with matrix elements $P(y|x) = \delta(y, f(x))$ for all $x \in S_{\underline{x}}$ and $y \in S_{\underline{y}}$. We will often talk about a map f and its associated probability matrix $P(y|x)$ as if they were the same thing.

Given two matrices A and B of the same dimensions, their **Hadamard product** $C = A \odot B$ is defined by $C_{i,j} = A_{i,j} B_{i,j}$ for all i, j . We will call $\text{HAS}(A) = A \odot A^\dagger$ the **Hadamard Absolute Square (HAS) of matrix A** . If U is a unitary matrix, then $\text{HAS}(U)$ is a probability matrix. For example, for any angle θ ,

$$\text{HAS}\left(\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}\right) = \begin{bmatrix} \cos^2 \theta & \sin^2 \theta \\ \sin^2 \theta & \cos^2 \theta \end{bmatrix}. \quad (101)$$

Another example is

$$\text{HAS}(\hat{H}_1) = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \quad (102)$$

A CB net \mathcal{N}^C is the **HAS of QB net \mathcal{N}^Q** if \mathcal{N}^Q and \mathcal{N}^C have the same graph, and their node matrices are related as follows. For each node \underline{x}_i , if $A[x_i|(x.)_{\Gamma_i}]$ is the amplitude of node \underline{x}_i in \mathcal{N}^Q , and $P[x_i|(x.)_{\Gamma_i}]$ is the probability of node \underline{x}_i in \mathcal{N}^C , then $|A[x_i|(x.)_{\Gamma_i}]|^2 = P[x_i|(x.)_{\Gamma_i}]$. In such a case, we will write $\text{HAS}(\mathcal{N}^Q) = \mathcal{N}^C$.

A unitary matrix $A(y, \tilde{x}|x, \tilde{y})$ (with rows labelled by y, \tilde{x} and columns by x, \tilde{y}) is a **q-embedding of probability matrix $P(y|x)$** if

$$\sum_{\tilde{x}} |A(y, \tilde{x}|x, \tilde{y} = 0)|^2 = P(y|x) \quad (103)$$

for all possible values of y and x . (the “q” in “q-embedding” stands for “quantum”). We say \tilde{y} is a **source index** and \tilde{x} is a **sink index**. We also refer to \tilde{x} and \tilde{y} collectively as **ancilla indices**. Note that any unitary matrix is a q-embedding of its HAS. Indeed, in this case Eq.(103) is satisfied with the indices \tilde{x} and \tilde{y} each ranging over a single value (i.e., \tilde{x} and \tilde{y} are fixed). If a q-embedding satisfies $A(y, \tilde{x}|x, \tilde{y}) \in \text{Bool}$ for all $y, \tilde{x}, x, \tilde{y}$, we say that it is a **deterministic q-embedding or a deterministic reversible extension (DRE) of its probability matrix** (note that its probability matrix must also be deterministic). By an extension of a matrix we mean adding extra rows and/or columns to it. General q-embeddings use the square root of the entries of the original probability matrix so they are not simply extensions of the original matrix; they are, however, reversible since they are unitary matrices.

Given a QB net \mathcal{N}^Q , let

$$P[(x.)_L] = \left| \sum_{(x.)_{\Gamma_Q - L}} A(x.) \right|^2. \quad (104)$$

On the right hand side of Eq.(104), $A(x.)$ is the amplitude of story $(x.)$, Γ_Q is the set of indices of all the nodes of \mathcal{N}^Q , and L is the set of indices of all leaf (aka external)

nodes of \mathcal{N}^Q . We say \mathcal{N}^Q is a **q-embedding of CB net** \mathcal{N}^C if $P[(x.)_L]$ defined by Eq.(104) satisfies

$$P[(x.)_{\Gamma_C}] = \sum_{L_1} P[(x.)_{L_1}] , \quad (105)$$

where $L_1 \subset L$, and Γ_C is the set of indices of all nodes of \mathcal{N}^C . Thus, the probability distribution associated with all nodes of \mathcal{N}^C can be obtained from the probability distribution associated with the external nodes of \mathcal{N}^Q . Some examples of q-embeddings of CB nets have already been given during our discussion of standard algorithms. More examples will be given in subsequent sections.

For some positive integers r and s , we will say a map $f : Bool^r \rightarrow Bool^s$ is a **binary gate** from r to s bits. f uniquely specifies (and is uniquely specified) by the deterministic probability matrix with entries $P(y|x) = \delta(f(x), y)$, where $x = (x_0, x_1, \dots, x_{r-1}) \in Bool^r$ and $y = (y_0, y_1, \dots, y_{s-1}) \in Bool^s$. If f is an invertible map, we will say that the gate is **reversible**. For example, the AND gate which takes $(x_1, x_0) \rightarrow y_0$ with $y_0 = x_0 x_1$ is a binary gate. So are the OR and NOT gates. Out of these 3 gates, only the NOT gate is reversible.

Another example of a reversible binary gate is the **Toffoli gate**[6]. It maps 3 bits into 3 bits as follows:

$$\begin{aligned} y_0 &= T_0(x) = x_0, \\ y_1 &= T_1(x) = x_1, \\ y_2 &= T_2(x) = x_2 \oplus x_0 x_1. \end{aligned} \quad (106)$$

The Toffoli gate can also be defined as the following deterministic probability matrix

$$P(y|x) = \delta(y, T(x)) = \delta(y_2, x_2 \oplus x_0 x_1) \delta(y_1, x_1) \delta(y_0, x_0) . \quad (107)$$

Consider 3 bits labelled 0, 1, and 2, and suppose the i th bit changes value from x_i to y_i . Then bits 0 and 1 do not change whereas bit 2 flips iff the product $x_0 x_1$ equals one. Thus, the probability matrix with entries given by Eq.(107) is simply a doubly controlled not:

$$[P(y|x)] = \sigma_x(2)^{n(1)n(0)} . \quad (108)$$

It is convenient to use the term Toffoli gate to refer not only to the gate defined by Eq.(107), but also to the 3 other gates that one obtains by replacing $x_0 x_1$ in Eq.(107) by $x_0 \bar{x}_1$, or $\bar{x}_0 x_1$, or $\bar{x}_0 \bar{x}_1$. This corresponds to replacing $n(1)n(0)$ in Eq.(108) by $n(1)\bar{n}(0)$, or $\bar{n}(1)n(0)$, or $\bar{n}(1)\bar{n}(0)$. Fig.8 shows the 4 doubly-controlled nots that we call Toffoli gates as well as the circuit diagrams usually used to represent them.

5.1 Q-Embedding of Probability Matrices

In this section we will first give some examples of q-embeddings of probability matrices. Then we will show that any probability matrix has a q-embedding.

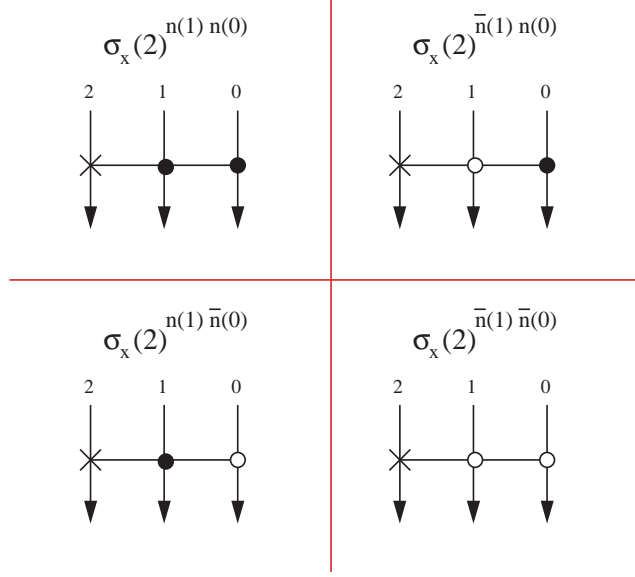


Figure 8: Four different kinds of Toffoli gates. 0,1,2 are bit labels.

Any unitary matrix is a q-embedding of its HAS, but such q-embeddings are trivial in the sense that they have no ancilla indices.

As first shown in Refs.[6], the Toffoli gates can be used to build q-embeddings (in fact, DREs) of the elementary binary gates AND, XOR, NOT, FANOUT. See Fig.9. Let $x = (x_0, x_1, x_2) \in Bool^3$ and $y = (y_0, y_1, y_2) \in Bool^3$. For the AND gate,

$$\sum_{y_1, y_0} \left| \langle y | \sigma_x(2)^{n(1)n(0)} | x_2 = 0, x_1, x_0 \rangle \right|^2 = \delta(y_2, x_1 x_0) . \quad (109a)$$

For the FANOUT gate,

$$\sum_{y_1} \left| \langle y | \sigma_x(2)^{\bar{n}(1)n(0)} | x_2 = 0, x_1 = 0, x_0 \rangle \right|^2 = \delta(y_2, x_0) \delta(y_0, x_0) . \quad (109b)$$

For the XOR gate,

$$\sum_{y_1, y_0} \left| \langle y | \sigma_x(2)^{n(1)\bar{n}(0)} | x_2, x_1, x_0 = 0 \rangle \right|^2 = \delta(y_2, x_2 \oplus x_1) . \quad (109c)$$

For the NOT gate,

$$\sum_{y_1, y_0} \left| \langle y | \sigma_x(2)^{\bar{n}(1)\bar{n}(0)} | x_2, x_1 = 0, x_0 = 0 \rangle \right|^2 = \delta(y_2, x_2 \oplus 1) = \delta(y_2, \bar{x}_2) . \quad (109d)$$

Note that the NOT gate is just σ_x , which is a DRE of itself. Eq.(109d) gives a different DRE of σ_x . In the left hand side of Eqs.(109), the x_i indices that are set to

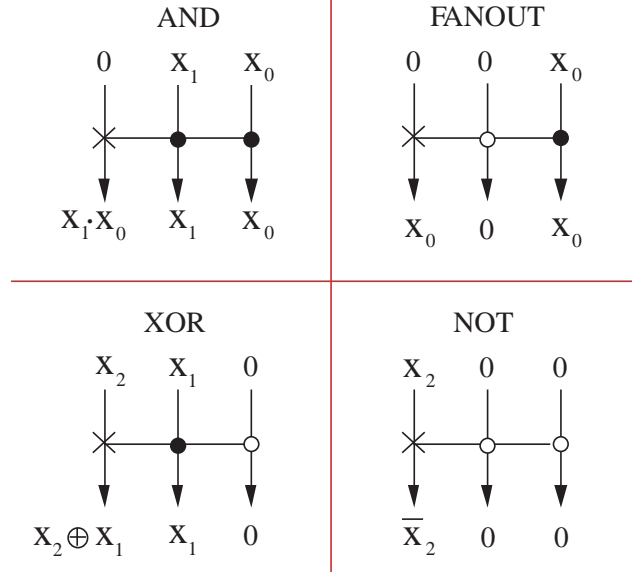


Figure 9: How to express elementary gates in terms of Toffoli gates. $x_0, x_1, x_2 \in Bool$ are bit values.

zero are called **source indices**, and the y_i indices that are summed over are called **sink indices**. Sink and source indices are collectively called **ancilla indices**.

Next we will prove that any probability matrix has a q-embedding. Suppose that we are given a probability matrix $P(y|x)$ where $x \in S_{\underline{x}}$ and $y \in S_{\underline{y}}$. Let $N_{\underline{x}}$ (ditto, $N_{\underline{y}}$) denote the number of elements in $S_{\underline{x}}$ (ditto, $S_{\underline{y}}$). Let $\xi^{(x)}$ for $x \in S_{\underline{x}}$ be any orthonormal basis of the complex $N_{\underline{x}}$ dimensional vector space. The components of $\xi^{(x)}$ will be denoted by $\xi_{\tilde{x}}^{(x)}$, where $\tilde{x} \in S_{\underline{x}}$. If the $\xi^{(x)}$'s are the standard basis, then $\xi_{\tilde{x}}^{(x)} = \delta(x, \tilde{x})$. Define matrix A by

$$A(y, \tilde{x}|x, \tilde{y}) = \begin{cases} \sqrt{P(y|x)} \xi_{\tilde{x}}^x & \text{if } \tilde{y} = 0 \\ \text{obtained by Gram-Schmidt method} & \text{if } \tilde{y} \neq 0 \end{cases} \quad (110)$$

To understand the last equation, consider Fig.10. In that figure we have assumed for definiteness that $S_{\underline{x}} = \{0, 1, 2\}$ and $S_{\underline{y}} = \{0, 1, 2, 3\}$. The shaded (ditto, unshaded) columns have $\tilde{y} \neq 0$ (ditto, $\tilde{y} = 0$). It is easy to see that the unshaded columns are orthonormal because the vectors ξ^x are orthonormal and $\sum_y P(y|x) = 1$. Since the unshaded columns are orthonormal, one can use the Gram-Schmidt method[10] to fill the shaded columns so that all the columns of A are orthonormal and therefore A is unitary. Note that by virtue of Eq.(110),

$$\sum_{\tilde{x}} |A(y, \tilde{x}|x, \tilde{y} = 0)|^2 = \sum_{\tilde{x}} P(y|x) \xi_{\tilde{x}}^{(x)*} \xi_{\tilde{x}}^{(x)} = P(y|x) \quad (111)$$

$y \quad \tilde{x}$
 \downarrow

$x \quad \tilde{y}$
 \longrightarrow

	00	01	02	03	10	11	12	13	20	21	22	23
00												
01	$\sqrt{P(0 0)}\xi^0$				$\sqrt{P(0 1)}\xi^1$				$\sqrt{P(0 2)}\xi^2$			
02												
10												
11	$\sqrt{P(1 0)}\xi^0$				$\sqrt{P(1 1)}\xi^1$				$\sqrt{P(1 2)}\xi^2$			
12												
20												
21	$\sqrt{P(2 0)}\xi^0$				$\sqrt{P(2 1)}\xi^1$				$\sqrt{P(2 2)}\xi^2$			
22												
30												
31	$\sqrt{P(3 0)}\xi^0$				$\sqrt{P(3 1)}\xi^1$				$\sqrt{P(3 2)}\xi^2$			
32												

Figure 10: How to construct a q-embedding of any probability matrix.

so that the A defined by Eq.(110) does indeed satisfy Eq.(103).[16]

Note that the matrix A defined by Eq.(110) has dimensions $N_{\underline{x}}N_{\underline{y}} \times N_{\underline{x}}N_{\underline{y}}$. It is sometimes possible to find a smaller q-embedding of an $N_{\underline{y}} \times N_{\underline{x}}$ probability matrix $P(y|x)$. For example, σ_x is a q-embedding of itself. As a less trivial example, suppose

$$P(y|x_1, x_2) = \delta(y, x_1 \oplus x_2) , \quad (112)$$

for $y, x_1, x_2 \in Bool$. Then define

$$A(y, e|x_1, x_2) = \frac{(-1)^{x_1 e}}{\sqrt{2}} \delta(y, x_1 \oplus x_2) , \quad (113)$$

for $y, e, x_1, x_2 \in Bool$. It is easy to check that matrix A is unitary. Furthermore,

$$\sum_e |A(y, e|x_1, x_2)|^2 = \frac{1}{2} \sum_e \delta(y, x_1 \oplus x_2) = \delta(y, x_1 \oplus x_2) . \quad (114)$$

5.2 Q-Embedding of CB Nets

As we've said before, F-T showed in Refs.[6] how, given any binary gate f , one can construct another binary gate \bar{f} such that \bar{f} is a DRE of f . Their method for constructing \bar{f} is to first represent f as a binary deterministic circuit composed of elementary gates (AND, XOR, NOT, FANOUT), and then to modify the circuit by replacing each of its gates by a DRE of it. The desired gate \bar{f} is then specified by the modified circuit.

In this section we will show how, given any CB net \mathcal{N}^C , one can construct a QB net \mathcal{N}^Q which is a q-embedding of \mathcal{N}^C . So far we've shown how to construct a q-embedding for any probability matrix. Now remember that each node of \mathcal{N}^C has a probability matrix assigned to it. The main step in constructing a q-embedding of \mathcal{N}^C is to replace each node matrix of \mathcal{N}^C with a q-embedding of it. Thus, our method for constructing a q-embedding of a CB net is a generalization of the F-T method for constructing a DRE of a binary deterministic circuit. We generalize their method so that it can be applied to any classical stochastic circuit, not just binary deterministic ones.

Before describing our construction method, we need some definitions. We say a node \underline{m} is a **marginalizer node** if it has a single input arrow and a single output arrow. Furthermore, the parent node of \underline{m} , call it \underline{x} , has states $x = (x_1, x_2, \dots, x_n)$, where $x_i \in S_{\underline{x}_i}$ for each $i \in Z_{1,n}$. Furthermore, for some particular integer $i_0 \in Z_{1,n}$, the set of possible states of \underline{m} is $S_{\underline{m}} = S_{\underline{x}_{i_0}}$, and the node matrix of \underline{m} is $P(\underline{m} = m | \underline{x} = x) = \delta(m, x_{i_0})$.

Let \mathcal{N}^C be a CB net for which we want to obtain a q-embedding. Our construction has two steps:

(Step 1) Add marginalizer nodes.

More specifically, replace \mathcal{N}^C by a modified CB net \mathcal{N}^C_{mod} obtained as follows. For each node \underline{x} of \mathcal{N}^C , add a marginalizer node between \underline{x} and every child of \underline{x} . If \underline{x} has no children, add a child to it.

As an example of this step, consider the net \mathcal{N}^C ("two body scattering net") defined by Fig.11 and Table 7.

nodes	states	probabilities	comments
\underline{a}	$a \in S_{\underline{a}}$	$P(a)$	
\underline{b}	$b \in S_{\underline{b}}$	$P(b)$	
\underline{c}	$c \in S_{\underline{c}}$	$P(c x)$	
\underline{d}	$d \in S_{\underline{d}}$	$P(d x)$	
\underline{x}	$x \in S_{\underline{x}}$	$P(x a, b)$	

Table 7

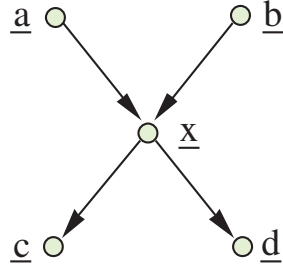


Figure 11: CB net for 2-body scattering. We show how to construct a q-embedding for this CB net.

Applying Step 1 to \mathcal{N}^C for two body scattering yields \mathcal{N}_{mod}^C defined by Fig.12 and Table 8.

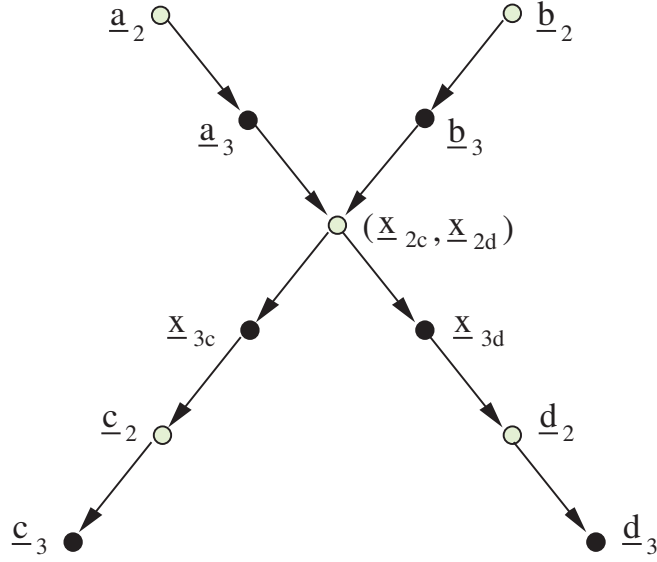


Figure 12: CB net of Fig.11 after adding marginalizer nodes.

nodes	states	probabilities	comments
\underline{a}_2	$a_2 \in S_{\underline{a}}$	$P_{\underline{a}}(a_2)$	
\underline{a}_3	$a_3 \in S_{\underline{a}}$	$\delta(a_3, a_2)$	
\underline{b}_2	$b_2 \in S_{\underline{b}}$	$P_{\underline{b}}(b_2)$	
\underline{b}_3	$b_3 \in S_{\underline{b}}$	$\delta(b_3, b_2)$	
\underline{c}_2	$c_2 \in S_{\underline{c}}$	$P_{\underline{c} \underline{x}}(c_2 x_{3c})$	
\underline{c}_3	$c_3 \in S_{\underline{c}}$	$\delta(c_3, c_2)$	
\underline{d}_2	$d_2 \in S_{\underline{d}}$	$P_{\underline{d} \underline{x}}(d_2 x_{3d})$	
\underline{d}_3	$d_3 \in S_{\underline{d}}$	$\delta(d_3, d_2)$	
$(\underline{x}_{2c}, \underline{x}_{2d})$	$(x_{2c}, x_{2d}) \in S_{\underline{x}}^2$	$P_{\underline{x} \underline{a}, \underline{b}}(x_{2c} a_3, b_3)\delta(x_{2d}, x_{2c})$	
\underline{x}_{3c}	$x_{3c} \in S_{\underline{x}}$	$\delta(x_{3c}, x_{2c})$	
\underline{x}_{3d}	$x_{3d} \in S_{\underline{x}}$	$\delta(x_{3d}, x_{2d})$	

Table 8

(Step 2) Replace node probability matrices by their q-embeddings. Add ancilla nodes.

More specifically, replace \mathcal{N}^C_{mod} by a QB net \mathcal{N}^Q obtained as follows. For each node of \mathcal{N}^C_{mod} , except for the marginalizer nodes that were added in the previous step, replace its node matrix by a new node matrix which is a q-embedding of the original node matrix. Add a new node for each ancilla index of each new node matrix. These new nodes will be called **ancilla nodes** (of either the source or sink type) because they correspond to ancilla indices.

Applying Step 2 to net \mathcal{N}^C_{mod} for two body scattering yields \mathcal{N}^Q defined by Fig.13 and Table 9.

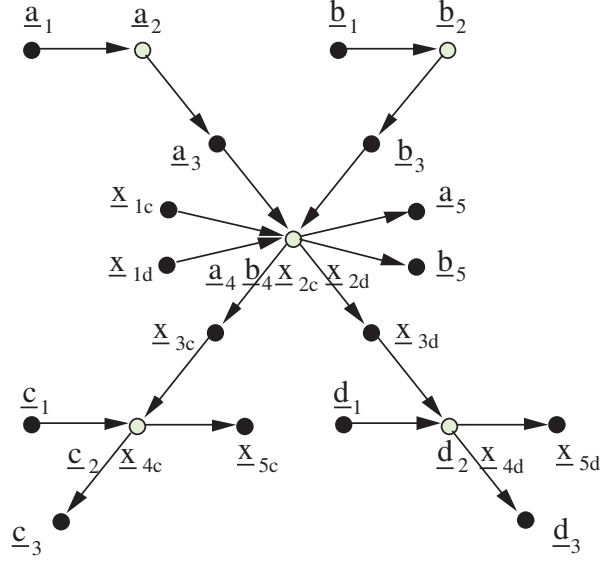


Figure 13: A QB Net which is a q-embedding for the CB net of Fig.11.

nodes	states	amplitudes	comments
\underline{a}_1	$a_1 \in S_{\underline{a}}$	$\delta(a_1, 0)$	
\underline{a}_2	$a_2 \in S_{\underline{a}}$	$A(a_2 a_1 = 0) = \sqrt{P_{\underline{a}}(a_2)}$	
\underline{a}_3	$a_3 \in S_{\underline{a}}$	$\delta(a_3, a_2)$	
$(\underline{a}_4, \underline{b}_4, \underline{x}_{2c}, \underline{x}_{2d})$	$(a_4, b_4, x_{2c}, x_{2d}) \in S_{\underline{a}, \underline{b}, \underline{x}, \underline{x}}$	$A(a_4, b_4, x_{2c}, x_{2d} a_3, b_3, x_{1c} = 0, x_{1d} = 0) = \sqrt{P_{\underline{x} \underline{a}, \underline{b}}(x_{2c} a_3, b_3)}\delta(a_4, a_3)\delta(b_4, b_3)\delta(x_{2d}, x_{2c})$	
\underline{a}_5	$a_5 \in S_{\underline{a}}$	$\delta(a_5, a_4)$	
\underline{b}_1	$b_1 \in S_{\underline{b}}$	$\delta(b_1, 0)$	
\underline{b}_2	$b_2 \in S_{\underline{b}}$	$A(b_2 b_1 = 0) = \sqrt{P_{\underline{b}}(b_2)}$	
\underline{b}_3	$b_3 \in S_{\underline{b}}$	$\delta(b_3, b_2)$	
\underline{b}_5	$b_5 \in S_{\underline{b}}$	$\delta(b_5, b_4)$	
\underline{c}_1	$c_1 \in S_{\underline{c}}$	$\delta(c_1, 0)$	
$(\underline{c}_2, \underline{x}_{4c})$	$(c_2, x_{4c}) \in S_{\underline{c}, \underline{x}}$	$A(c_2, x_{4c} c_1 = 0, x_{3c}) = \sqrt{P_{\underline{c} \underline{x}}(c_2 x_{3c})}\delta(x_{4c}, x_{3c})$	
\underline{c}_3	$c_3 \in S_{\underline{c}}$	$\delta(c_3, c_2)$	
\underline{d}_1	$d_1 \in S_{\underline{d}}$	$\delta(d_1, 0)$	
$(\underline{d}_2, \underline{x}_{4d})$	$(d_2, x_{4d}) \in S_{\underline{d}, \underline{x}}$	$A(d_2, x_{4d} d_1 = 0, x_{3d}) = \sqrt{P_{\underline{d} \underline{x}}(d_2 x_{3d})}\delta(x_{4d}, x_{3d})$	
\underline{d}_3	$d_3 \in S_{\underline{d}}$	$\delta(d_3, d_2)$	
\underline{x}_{1c}	$x_{1c} \in S_{\underline{x}}$	$\delta(x_{1c}, 0)$	
\underline{x}_{1d}	$x_{1d} \in S_{\underline{x}}$	$\delta(x_{1d}, 0)$	
\underline{x}_{3c}	$x_{3c} \in S_{\underline{x}}$	$\delta(x_{3c}, x_{2c})$	
\underline{x}_{3d}	$x_{3d} \in S_{\underline{x}}$	$\delta(x_{3d}, x_{2d})$	
\underline{x}_{5c}	$x_{5c} \in S_{\underline{x}}$	$\delta(x_{5c}, x_{4c})$	
\underline{x}_{5d}	$x_{5d} \in S_{\underline{x}}$	$\delta(x_{5d}, x_{4d})$	

Table 9

\mathcal{N}^Q looks much more complicated than \mathcal{N}^C , but it really isn't, since most of its node matrices are delta functions which quickly disappear when adding over node states.

According to Table 9, the probability amplitude for the external (aka leaf) nodes is given by

$$\begin{aligned}
& A(a_5, b_5, c_3, d_3, x_{5c}, x_{5d}) = \\
& = \sum. \sqrt{P_{\underline{a}}(a_2)P_{\underline{b}}(b_2)P_{\underline{x}|\underline{a},\underline{b}}(x_{2c}|a_3, b_3)P_{\underline{c}|\underline{x}}(c_2|x_{3c})P_{\underline{d}|\underline{x}}(d_2|x_{3d})} \\
& \quad \theta(a_2 = a_3 = a_4 = a_5)\theta(b_2 = b_3 = b_4 = b_5) \\
& \quad \theta(x_{2c} = x_{3c} = x_{4c} = x_{5c})\theta(x_{2d} = x_{3d} = x_{4d} = x_{5d})\theta(x_{5c} = x_{5d}) \\
& \quad \theta(c_2 = c_3)\theta(d_2 = d_3) \\
& \quad \theta(a_1 = b_1 = c_1 = d_1 = x_{1c} = x_{1d} = 0)
\end{aligned} \tag{115}$$

where we have summed over all internal (non-leaf) nodes. Eq.(115) immediately reduces to

$$\begin{aligned}
& A(a_5, b_5, c_3, d_3, x_{5c}, x_{5d}) = \\
& = \sqrt{P_{\underline{a}}(a_5)P_{\underline{b}}(b_5)P_{\underline{x}|\underline{a},\underline{b}}(x_{5c}|a_5, b_5)P_{\underline{c}|\underline{x}}(c_3|x_{5c})P_{\underline{d}|\underline{x}}(d_3|x_{5d})\theta(x_{5c} = x_{5d})}
\end{aligned} \tag{116}$$

Eq.(116) shows that the net \mathcal{N}^Q that we constructed from the net \mathcal{N}^C by following steps 1 and 2 satisfies the definition Eq.(105) that we gave earlier for a q-embedding of \mathcal{N}^C . The probability distribution of the states of the external nodes of the QB net \mathcal{N}^Q contains all the probabilistic information of the original CB net \mathcal{N}^C . Hurray!

From Eq.(116), it is clear that by running \mathcal{N}^Q on a quantum computer (or similar quantum system), we can calculate any conditional probability that one would want to calculate for \mathcal{N}^C . For example, suppose we wanted to calculate $P_{\underline{a},\underline{d}|\underline{x}}$. Run \mathcal{N}^Q on the quantum computer several times, each time measuring nodes $\underline{a}_5, \underline{d}_3$ and \underline{x}_{5d} and not measuring all other external nodes. The resulting measurements will be distributed according to $P_{\underline{a},\underline{d},\underline{x}}$. Taking the magnitude squared of the amplitude and summing the result over the states of the un-measured external nodes will be performed automatically by Nature. The laws of quantum mechanics guarantee it. Proceed in the same way to calculate $P_{\underline{x}}$. Run \mathcal{N}^Q on the quantum computer several times, each time measuring node \underline{x}_{5d} and not measuring all other external nodes. Finally divide $P_{\underline{a},\underline{d},\underline{x}}$ by $P_{\underline{x}}$ on a classical (or quantum?) computer.

The q-embedding of a CB net, as defined by Eq.(105), is not unique. For example, we could have defined the net \mathcal{N}^Q given by Fig.13 without nodes \underline{a}_3 and \underline{b}_3 . We chose to include such nodes for pedagogical reasons.

To run a QB net on a quantum computer, we need to replace the QB net by an equivalent SEO that a quantum computer can understand. This can be done with the help of a quantum compiler [9][8]. One could compile individually each node representing a q-embedding, or one could compile whole subgraphs of the QB

net all at once. Note that it may suffice to find a SEO that is only approximately (within a certain precision) equivalent instead of exactly equivalent to the QB net. This may be true if, for example, the probabilities associated with the CB net that was q-embedded were not specified too precisely to begin with.

Suppose a_1, a_2, \dots, a_ν belong to a finite set $S_{\underline{a}}$, and suppose that they are distributed according to a probability distribution $P_{\underline{a}}$. What number ν of samples a_i is necessary to estimate $P_{\underline{a}}$ within a given precision? This question is directly relevant to our method for estimating probabilities by running a QB net on a quantum computer. We will not give a detailed answer to this question here. For an answer, the reader can consult any book on the mathematical theory of Statistics. An imprecise rule of thumb is that if the support of $P_{\underline{a}}$ has ν_0 elements, then ν must be at least as large as ν_0 ; i.e., one needs at least “one data point per bin” to estimate $P_{\underline{a}}$ with any decent accuracy.

We have given a method for calculating, via a quantum computer, the conditional probabilities associated with a CB net. Does our method have an advantage in time complexity with respect to classical methods for calculating the same probabilities? We will not give a detailed answer to this question here. The answer must be yes, sometimes. After all, our method generalizes the algorithms by Deutsch-Jozsa, Simon, Grover, etc., and these are known to have a complexity advantage.

To conclude this section, we will present a second, more complicated example of our method of finding a q-embedding for a CB net. A CB net (first given in Ref.[17]) for lung disease diagnosis is defined by Fig.14 and Table 10.

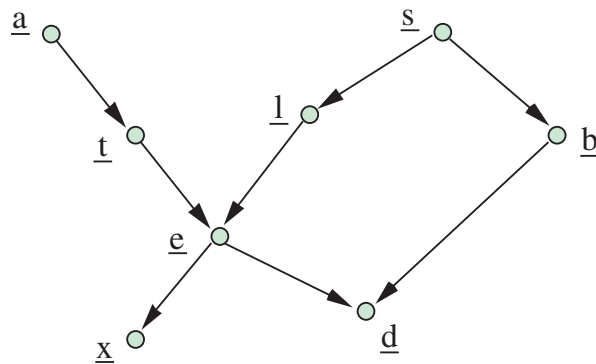


Figure 14: CB net for lung disease diagnosis. We show how to construct a q-embedding for this CB net.

nodes	states	probabilities	comments
<u>a</u>	$a \in Bool$	$P(\underline{a} = 1) = .01$	Visited Asia?
<u>b</u>	$b \in Bool$	$P(\underline{b} = 1 \underline{s} = 1) = .60$ $P(\underline{b} = 1 \underline{s} = 0) = .30$	Bronchitis?
<u>d</u>	$d \in Bool$	$P(\underline{d} = 1 \underline{e} = 1, \underline{b} = 1) = .90$ $P(\underline{d} = 1 \underline{e} = 1, \underline{b} = 0) = .70$ $P(\underline{d} = 1 \underline{e} = 0, \underline{b} = 1) = .80$ $P(\underline{d} = 1 \underline{e} = 0, \underline{b} = 0) = .10$	Dyspnea(trouble breathing)?
<u>e</u>	$e \in Bool$	$P(e l, t) = \delta(e, l \vee t)$	Either TB or Lung Cancer?
<u>l</u>	$l \in Bool$	$P(\underline{l} = 1 \underline{s} = 1) = .10$ $P(\underline{l} = 1 \underline{s} = 0) = .01$	Lung Cancer?
<u>s</u>	$s \in Bool$	$P(\underline{s} = 1) = .5$	Smokes?
<u>t</u>	$t \in Bool$	$P(\underline{t} = 1 \underline{a} = 1) = .05$ $P(\underline{t} = 1 \underline{a} = 0) = .01$	Tuberculosis?
<u>x</u>	$x \in Bool$	$P(\underline{x} = 1 \underline{e} = 1) = .98$ $P(\underline{x} = 1 \underline{e} = 0) = .05$	Positive X-ray?

Table 10

If one follows the two steps that were described earlier in this section, one obtains the QB net defined by Fig.15 and Table 11.

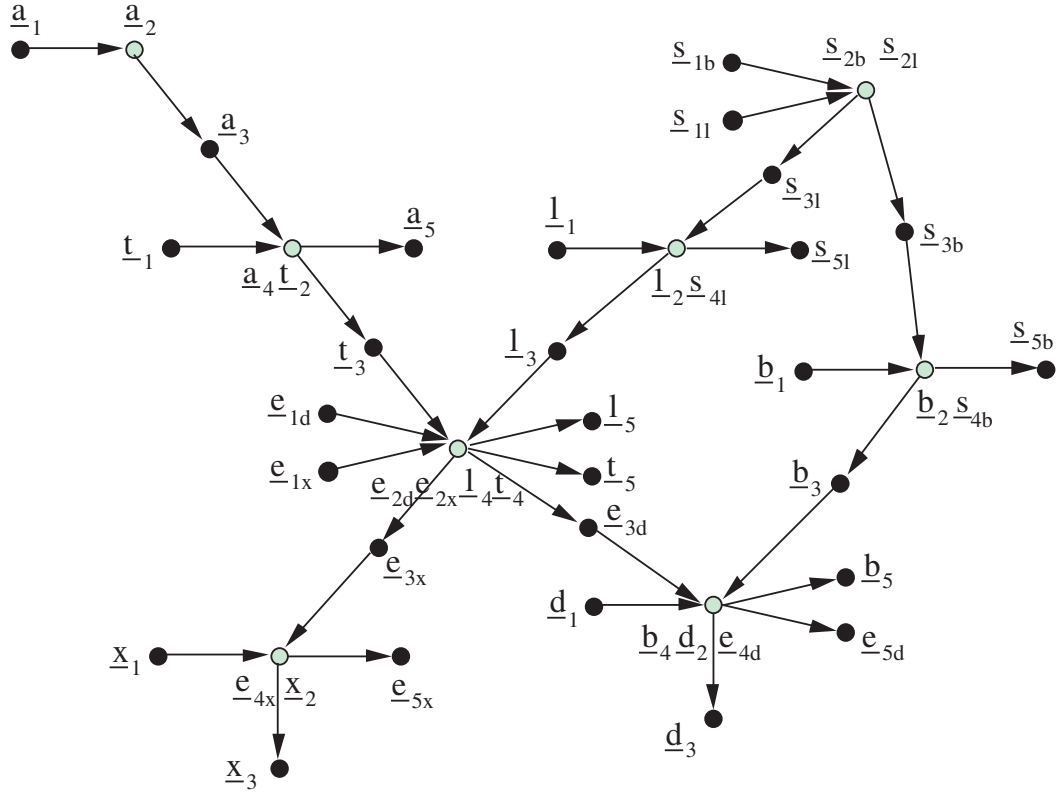


Figure 15: A QB Net which is a q-embedding for the CB net of Fig.14.

nodes	states	amplitudes
\underline{a}_1	$a_1 \in Bool$	$\delta(a_1, 0)$
\underline{a}_2	$a_2 \in Bool$	$A(a_2 a_1 = 0) = \sqrt{P_{\underline{a}}(a_2)}$
\underline{a}_3	$a_3 \in Bool$	$\delta(a_3, a_2)$
$(\underline{a}_4, \underline{t}_2)$	$(a_4, t_2) \in Bool^2$	$A(a_4, t_2 a_3, t_1 = 0) = \sqrt{P_{\underline{t} \underline{a}}(t_2 a_3)}\delta(a_4, a_3)$
\underline{a}_5	$a_5 \in Bool$	$\delta(a_5, a_4)$
\underline{b}_1	$b_1 \in Bool$	$\delta(b_1, 0)$
$(\underline{b}_2, \underline{s}_{4b})$	$(b_2, s_{4b}) \in Bool^2$	$A(b_2, s_{4b} b_1 = 0, s_{3b}) = \sqrt{P_{\underline{b} \underline{s}}(b_2 s_{3b})}\delta(s_{4b}, s_{3b})$
\underline{b}_3	$b_3 \in Bool$	$\delta(b_3, b_2)$
$(\underline{b}_4, \underline{d}_2, \underline{e}_{4d})$	$(b_4, d_2, e_{4d}) \in Bool^3$	$A(b_4, d_2, e_{4d} b_3, d_1 = 0, e_{3d}) = \sqrt{P_{\underline{d} \underline{b}, \underline{e}}(d_2 b_3, e_{3d})}\delta(b_4, b_3)\delta(e_{4d}, e_{3d})$
\underline{b}_5	$b_5 \in Bool$	$\delta(b_5, b_4)$
\underline{d}_1	$d_1 \in Bool$	$\delta(d_1, 0)$
\underline{d}_3	$d_3 \in Bool$	$\delta(d_3, d_2)$
\underline{e}_{1d}	$e_{1d} \in Bool$	$\delta(e_{1d}, 0)$
\underline{e}_{1x}	$e_{1x} \in Bool$	$\delta(e_{1x}, 0)$
$(\underline{e}_{2d}, \underline{e}_{2x}, \underline{l}_4, \underline{t}_4)$	$(e_{2d}, e_{2x}, l_4, t_4) \in Bool^4$	$A(e_{2d}, e_{2x}, l_4, t_4 e_{1d} = 0, e_{1x} = 0, l_3, t_3) = \sqrt{P_{\underline{e} \underline{l}, \underline{t}}(e_{2d} l_3, t_3)}\delta(e_{2x}, e_{2d})\delta(l_4, l_3)\delta(t_4, t_3)$
\underline{e}_{3d}	$e_{3d} \in Bool$	$\delta(e_{3d}, e_{2d})$
\underline{e}_{3x}	$e_{3x} \in Bool$	$\delta(e_{3x}, e_{2x})$
$(\underline{e}_{4x}, \underline{x}_2)$	$(e_{4x}, x_2) \in Bool^2$	$A(e_{4x}, x_2 e_{3x}, x_1 = 0) = \sqrt{P_{\underline{x} \underline{e}}(x_2 e_{3x})}\delta(e_{4x}, e_{3x})$
\underline{e}_{5d}	$e_{5d} \in Bool$	$\delta(e_{5d}, e_{4d})$
\underline{e}_{5x}	$e_{5x} \in Bool$	$\delta(e_{5x}, e_{4x})$
\underline{l}_1	$l_1 \in Bool$	$\delta(l_1, 0)$
$(\underline{l}_2, \underline{s}_{4l})$	$(l_2, s_{4l}) \in Bool^2$	$A(l_2, s_{4l} l_1 = 0, s_{3l}) = \sqrt{P_{\underline{l} \underline{s}}(l_2 s_{3l})}\delta(s_{4l}, s_{3l})$
\underline{l}_3	$l_3 \in Bool$	$\delta(l_3, l_2)$
\underline{l}_5	$l_5 \in Bool$	$\delta(l_5, l_4)$
\underline{s}_{1b}	$s_{1b} \in Bool$	$\delta(s_{1b}, 0)$
\underline{s}_{1l}	$s_{1l} \in Bool$	$\delta(s_{1l}, 0)$
$(\underline{s}_{2b}, \underline{s}_{2l})$	$(s_{2b}, s_{2l}) \in Bool^2$	$A(s_{2b}, s_{2l} s_{1b} = 0, s_{1l} = 0) = \sqrt{P_{\underline{s}}(s_{2b})}\delta(s_{2l}, s_{2b})$
\underline{s}_{3b}	$s_{3b} \in Bool$	$\delta(s_{3b}, e_{2b})$
\underline{s}_{3l}	$s_{3l} \in Bool$	$\delta(s_{3l}, s_{2l})$
\underline{s}_{5b}	$s_{5b} \in Bool$	$\delta(s_{5b}, s_{4b})$
\underline{s}_{5l}	$s_{5l} \in Bool$	$\delta(s_{5l}, s_{4l})$
\underline{t}_1	$t_1 \in Bool$	$\delta(t_1, 0)$
\underline{t}_3	$t_3 \in Bool$	$\delta(t_3, t_2)$
\underline{t}_5	$t_5 \in Bool$	$\delta(t_5, t_4)$
\underline{x}_1	$x_1 \in Bool$	$\delta(x_1, 0)$
\underline{x}_3	$x_3 \in Bool$	$\delta(x_3, x_2)$

Table 11

According to Table 11, the probability amplitude for the external (aka leaf) nodes is given by

$$\begin{aligned}
& A(a_5, b_5, d_3, e_{5d}, e_{5x}, l_5, s_{5b}, s_{5l}, t_5, x_3) = \\
& = \sqrt{\frac{P_{\underline{a}}(a_5)P_{\underline{t}|\underline{a}}(t_5|a_5)P_{\underline{b}|\underline{s}}(b_5|s_{5b})P_{\underline{d}|\underline{b},\underline{e}}(d_3|b_5, e_{5d})P_{\underline{e}|\underline{l},\underline{t}}(e_{5d}|l_5, t_5)P_{\underline{x}|\underline{e}}(x_3|e_{5d})P_{\underline{l}|\underline{s}}(l_5|s_{5l})P_{\underline{s}}(s_{5b})}{\theta(e_{5d}, e_{5x})\theta(s_{5b}, s_{5l})}} \quad .
\end{aligned} \tag{117}$$

6 Voting Net and Grover's Microscope

In this section we will first present a CB net, call it \mathcal{N}^C , that describes voting. Then we will find a QB net \mathcal{N}^Q that is a q-embedding of \mathcal{N}^C . In certain cases, the probabilities that we wish to find are too small to be measurable by running \mathcal{N}^Q on a quantum computer. However, we will show that sometimes it is possible to define a new QB net, call it $\mathcal{N}^{Q'}$, that magnifies and makes measurable the probabilities that were unmeasurable using \mathcal{N}^Q alone. We will refer to $\mathcal{N}^{Q'}$ as **Grover's Microscope for \mathcal{N}^Q** , because $\mathcal{N}^{Q'}$ is closely related to Grover's algorithm, and it magnifies the probabilities found with \mathcal{N}^Q .

Suppose $y \in Bool$ and $\vec{x} = (x^0, x^1, \dots, x^{N_B-1}) \in Bool^{N_B}$. Consider the CB net ("voting net") defined by Fig.16 and Table 12.

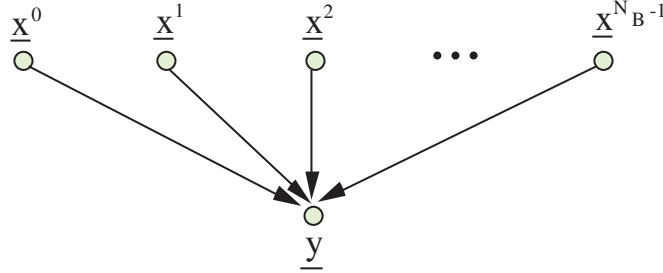


Figure 16: "Voting" CB net.

nodes	states	probabilities	comments
\underline{x}^i for all $i \in Z_{0,N_B-1}$	$x^i \in Bool$	$P(x^i)$	
\underline{y}	$y \in Bool$	$P(y \vec{x})$	

Table 12

Henceforth, we will abbreviate $P(y = 0|\vec{x}) = p_i$ and $P(y = 1|\vec{x}) = q_i$, where $i = dec(\vec{x}) \in Z_{0,N_S-1}$. Hence $p_i + q_i = 1$ for all $i \in Z_{0,N_S-1}$. In general, the probability

matrix $P(y|\vec{x})$ has 2^{N_B} free parameters (namely, p_i for all $i \in Z_{0, N_B-1}$). This number of parameters is a forbiddingly large for large N_B . To ease the task of specifying $P(y|\vec{x})$, it is common to impose additional constraints on $P(y|\vec{x})$. An interesting special type of $P(y|\vec{x})$ is **deterministic** $pd(Bool|Bool^{N_B})$ **matrices**; that is, those that can be expressed in the form

$$P(y|\vec{x}) = \delta(y, f(\vec{x})) , \quad (118)$$

where $f : Bool^{N_B} \rightarrow Bool$. In this case, the voting net can be used to pose the **satisfiability problem (SAT)**: given $y = 0$, find the most likely $\vec{x} \in Bool^{N_B}$; in other words, find those \vec{x} for which $f(\vec{x}) = 0$. We say f is **AND-like** if all p_i equal zero except for one p_i which equals one. For example, for $N_B = 2$, if f is an AND gate, then

$$P(y|\vec{x})_{AND} = \left\{ \begin{array}{c|cccc} & & \begin{matrix} (x^0, x^1) \rightarrow \\ 00 \quad 01 \quad 10 \quad 11 \end{matrix} \\ \hline y \downarrow & 0 & 1 & 1 & 1 & 0 \\ & 1 & 0 & 0 & 0 & 1 \end{array} \right. . \quad (119)$$

A slightly more general type of $P(y|\vec{x})$ is **quasi-deterministic** $pd(Bool|Bool^{N_B})$ **matrices**; that is, those that can be expressed in the form

$$P(y|\vec{x}) = \sum_{\vec{t}} \delta(y, f(\vec{t})) P(t^0|x^0) P(t^1|x^1) \dots P(t^{N_B-1}|x^{N_B-1}) , \quad (120)$$

where $f : Bool^{N_B} \rightarrow Bool$ and we sum over all $\vec{t} = (t^0, t^1, \dots, t^{N_B-1}) \in Bool^{N_B}$. When $f(\vec{t}) = t^0 \vee t^1 \vee \dots \vee t^{N_B-1}$, $P(y|\vec{x})$ is called a **noisy-OR**. Appendix A discusses how to q-embed deterministic $pd(Bool|Bool^{N_B})$ matrices, and how to express such q-embeddings as a SEO. Appendix B discusses the same thing for quasi-deterministic $pd(Bool|Bool^{N_B})$ matrices.

A q-embedding for the CB net defined by Fig.16 and Table 12 is given by the QB net defined by Fig.17 and Table 13.

nodes	states	amplitudes	comments
$\underline{\vec{x}}_1$	$\vec{x}_1 \in Bool^{N_B}$	$\delta(\vec{x}_1, 0)$	
$\underline{\vec{x}}_2$	$\vec{x}_2 \in Bool^{N_B}$	$A(\vec{x}_2 \vec{x}_1 = 0) = \sqrt{P_{\underline{\vec{x}}}(\vec{x}_2)}$	
$(\underline{\vec{x}}_3, \underline{y}_2)$	$(\vec{x}_3, y_2) \in Bool^{N_B+1}$	$A(\vec{x}_3, y_2 \vec{x}_2, y_1 = 0) = \sqrt{P_{\underline{y} \underline{\vec{x}}}(y_2 \vec{x}_2)} \delta(\vec{x}_3, \vec{x}_2)$	
$\underline{\vec{x}}_4$	$\vec{x}_4 \in Bool^{N_B}$	$\delta(\vec{x}_4, \vec{x}_3)$	
\underline{y}_1	$y_1 \in Bool$	$\delta(y_1, 0)$	
\underline{y}_3	$y_3 \in Bool$	$\delta(y_3, y_2)$	

Table 13

According to Table 13, the probability amplitude for the leaf (external) nodes is

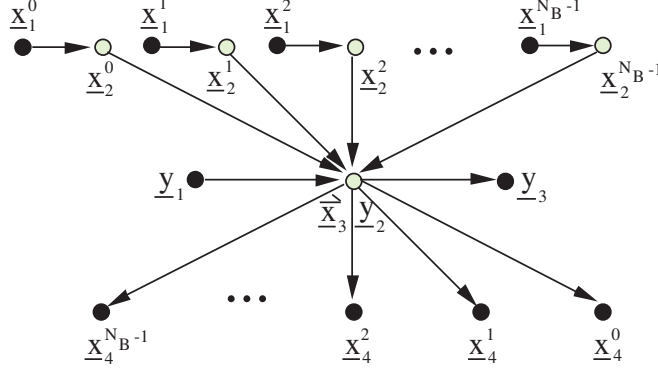


Figure 17: A QB Net which is a q-embedding for the CB net of Fig.16.

$$A(\vec{x}_4, y_3) = \sum \sqrt{P_{\underline{x}}(\vec{x}_2)P_{y|\underline{x}}(y_2|\vec{x}_2)}\theta(y_2 = y_3)\theta(\vec{x}_2 = \vec{x}_3 = \vec{x}_4)\theta(\vec{x}_1 = y_1 = 0) \quad (121a)$$

$$= \sqrt{P_{\underline{x}}(\vec{x}_4)P_{y|\underline{x}}(y_3|\vec{x}_4)} . \quad (121b)$$

To fully specify the QB net for voting, we need to extend $A(\vec{x}_2|\vec{x}_1 = 0)$ and $A(\vec{x}_3, y_2|\vec{x}_2, y_1 = 0)$ into unitary matrices by adding columns to them. This can always be accomplished by applying the Gram-Schmidt algorithm. But sometimes one can guess a matrix extension and applying Gram-Schmidt becomes unnecessary. If $P_{\underline{x}}$ is uniform (i.e., $P(\vec{x}) = 1/N_S$ for all \vec{x} , which means there is no a priori information about \vec{x}), then $A(\vec{x}_2|\vec{x}_1 = 0) = 1/\sqrt{N_S}$. In this case, we can extend $A(\vec{x}_2|\vec{x}_1 = 0)$ into the unitary matrix

$$[A(\vec{x}_2|\vec{x}_1)] = \hat{H}_{N_B} . \quad (122)$$

(This works because all entries of the first column of \hat{H}_{N_B} are equal to $1/\sqrt{N_S}$.) As to extending $A(\vec{x}_3, y_2|\vec{x}_2, y_1 = 0)$, this can be done as follows. Define

$$\Delta_p = \text{diag}(\sqrt{p_0}, \sqrt{p_1}, \dots, \sqrt{p_{N_S-1}}) , \quad (123)$$

and

$$\Delta_q = \text{diag}(\sqrt{q_0}, \sqrt{q_1}, \dots, \sqrt{q_{N_S-1}}) . \quad (124)$$

A possible way of extending $A(\vec{x}_3, y_2|\vec{x}_2, y_1 = 0)$ into a unitary matrix is

$$[A(\vec{x}_3, y_2|\vec{x}_2, y_1)] = \begin{pmatrix} \Delta_p & -\Delta_q \\ \Delta_q & \Delta_p \end{pmatrix} . \quad (125)$$

Unitary matrices of this kind are called D-matrices in Ref.[8]. Ref.[8] shows how to decompose any D-matrix into a SEO.

Earlier, we explained how to estimate a conditional probability for a CB net by running a QB net ν times on a quantum computer. If we wanted to find $P(y|x^0, x^1)$ for the voting CB net, then the number of runs ν required to estimate $P(y|x^0, x^1)$ with moderate accuracy would not be too onerous, because the domain of $P(y|x^0, x^1)$ is $Bool^3$, which contains only 8 points. But what if we wanted to estimate $P(y|\vec{x})$? For large N_B , the domain of $P(y|\vec{x})$ is very large (2^{N_B+1} points). If the support of $P(y|\vec{x})$ occupies a large fraction of this domain, then the number of runs ν required to estimate $P(y|\vec{x})$ with moderate accuracy is forbiddingly large. However, there are some cases in which ‘‘Grover’s Microscope’’ can come to the rescue, by allowing us to amplify certain salient features of $P(y|\vec{x})$ so that they become measurable in only a few runs.

Next we will discuss Grover’s Microscope for the voting QB net defined by Fig.17 and Table 13. For simplicity, we will assume that $P_{\vec{x}}$ is uniform.

Let $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$ label N_B bits and let τ label another bit. Assume that τ and all the κ_i are distinct. Define

$$|\phi_p\rangle_{\vec{\kappa}} = \phi_p = (\sqrt{p_0}, \sqrt{p_1}, \dots, \sqrt{p_{N_S-1}})^T, \quad (126)$$

$$|\phi_q\rangle_{\vec{\kappa}} = \phi_q = (\sqrt{q_0}, \sqrt{q_1}, \dots, \sqrt{q_{N_S-1}})^T, \quad (127)$$

and

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{N_S}} (|0\rangle_{\tau} |\phi_p\rangle_{\vec{\kappa}} + |1\rangle_{\tau} |\phi_q\rangle_{\vec{\kappa}}) \\ &= \frac{1}{\sqrt{N_S}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \phi_p + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \phi_q \right] = \frac{1}{\sqrt{N_S}} \begin{pmatrix} \phi_p \\ \phi_q \end{pmatrix} = \Psi. \end{aligned} \quad (128)$$

Since $p_i + q_i = 1$ for all i , $\phi_p^T \phi_p + \phi_q^T \phi_q = N_S$. According to Eq.(121), when $P_{\vec{x}}$ is uniform, the voting QB net fully specifies a unitary matrix U_{net} such that

$$|\Psi\rangle = U_{net} |0\rangle_{\vec{\kappa}} |0\rangle_{\tau}. \quad (129)$$

Define orthonormal vectors e_0 and e_1 by

$$e_0 = \begin{pmatrix} \hat{\phi}_p \\ 0 \end{pmatrix}, \quad e_1 = \begin{pmatrix} 0 \\ \hat{\phi}_q \end{pmatrix}, \quad (130)$$

where \hat{V} is a unit vector in the direction of \vec{V} . If $P(y|\vec{x})$ is deterministic with AND-like f , then all components of e_0 are zero except for the one at the target state j_{targ} .

In terms of e_0, e_1 , Ψ can be expressed as

$$\Psi = \frac{1}{\sqrt{N_S}} \begin{pmatrix} \phi_p \\ \phi_q \end{pmatrix} = \frac{1}{\sqrt{N_S}} (|\phi_p|e_0 + |\phi_q|e_1) . \quad (131)$$

It is convenient to define a vector Ψ_\perp orthogonal to Ψ :

$$\Psi_\perp = \frac{1}{\sqrt{N_S}} (|\phi_q|e_0 - |\phi_p|e_1) . \quad (132)$$

If $P(y|\vec{x})$ is deterministic with AND-like f , then $|\phi_p| = 1$ and $|\phi_q| = \sqrt{N_S - 1}$ so, for large N_S , $\Psi \approx e_1$ and $\Psi_\perp \approx e_0$. For an arbitrary angle α , let

$$\Psi'_\perp = \frac{1}{\sqrt{N_S}} \left[(c_{\frac{\alpha}{2}}|\phi_q| + s_{\frac{\alpha}{2}}|\phi_p|)e_0 + (s_{\frac{\alpha}{2}}|\phi_q| - c_{\frac{\alpha}{2}}|\phi_p|)e_1 \right] , \quad (133)$$

where $s_A = \sin A$ and $c_A = \cos A$ for any angle A . Let $\angle(x, y)$ denote the angle between 2 vectors x and y . Note that $\angle(\Psi'_\perp, \Psi_\perp) = \alpha/2$. We define $\angle(e_1, \Psi) = \theta/2$.

Fig.18 portrays various vectors that arise in explaining Grover's Microscope. Note that $\Psi'_\perp = e_0$ when $\alpha = \theta$.

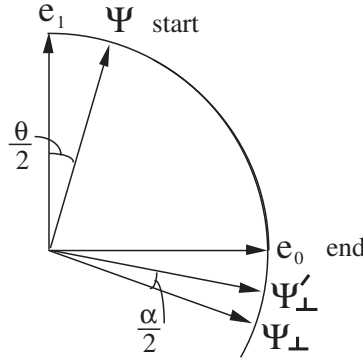


Figure 18: Various vectors relevant to Grover's Microscope.

Since we plan to stay within the two dimensional vector space with orthonormal basis e_0, e_1 , it is convenient to switch matrix representations. Within $\text{span}(e_0, e_1)$, e_0, e_1 can be represented more simply by:

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (134)$$

If e_0, e_1 are represented in this way, then

$$\Psi = \frac{1}{\sqrt{N_S}} \begin{pmatrix} |\phi_p| \\ |\phi_q| \end{pmatrix} , \quad (135)$$

$$\Psi_\perp = \frac{1}{\sqrt{N_S}} \begin{pmatrix} |\phi_q| \\ -|\phi_p| \end{pmatrix} , \quad (136)$$

and

$$\Psi'_\perp = W\Psi, \text{ where } W = \begin{pmatrix} c_{\frac{\alpha}{2}} & -s_{\frac{\alpha}{2}} \\ s_{\frac{\alpha}{2}} & c_{\frac{\alpha}{2}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (137)$$

The matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is a clockwise rotation by $\pi/2$ in space $\text{span}(e_0, e_1)$. Thus, W equals a clockwise rotation by $\pi/2$ followed by a counter-clockwise rotation by $\alpha/2$.

Define the following reflection operators

$$R_0 = 1 - 2\Pi_{|0\rangle_{\bar{\kappa}}}\Pi_{|0\rangle_{\tau}} = (-1)^{\Pi_{|0\rangle_{\bar{\kappa}}}\Pi_{|0\rangle_{\tau}}}, \quad (138)$$

$$R_\Psi = U_{net}R_0U_{net}^\dagger, \quad (139)$$

$$R_{\Psi'_\perp} = WR_\Psi W^\dagger = WU_{net}R_0U_{net}^\dagger W^\dagger. \quad (140)$$

From Eq.(24), it follows that

$$-R_\Psi R_{\Psi'_\perp} = c_\alpha \Psi \Psi^T - s_\alpha \Psi \Psi_\perp^T + s_\alpha \Psi_\perp \Psi^T + c_\alpha \Psi_\perp \Psi_\perp^T. \quad (141)$$

Thus, $-R_\Psi R_{\Psi'_\perp}$ rotates vectors in $\text{span}(e_0, e_1)$, clockwise by an angle α .

Grover's Microscope can be summarized by the following equation

$$(-R_\Psi R_{\Psi'_\perp})^r \Psi \approx e_0, \quad (142)$$

for some integer r to be determined, where “ \approx ” means approximation at large N_S . What this means is that our system starts in state Ψ and is rotated consecutively r times, each time by a small angle α , until it arrives at the state e_0 . If $P(y|\vec{x})$ is deterministic with AND-like f , then measuring state e_0 yields the target state j_{targ} .

The optimum number r of iterations is

$$r\alpha \approx \frac{\pi}{2}(1 + 2k) \quad (143)$$

for some integer k . Note that $\cos(\theta/2) = \langle \Psi | e_1 \rangle = |\phi_q|/\sqrt{N_S}$ so, in general, θ depends on $|\phi_p|$ (or on $|\phi_q| = \sqrt{N_S - |\phi_p|^2}$). If $P(y|\vec{x})$ is deterministic with AND-like f , then $|\phi_p| = 1$ and $|\phi_q| = \sqrt{N_S - 1}$. In this case, it is convenient to choose $\alpha = \theta$, so that $\Psi'_\perp = e_1$ and Figs.6 and 18 become the same diagram under the mapping $\Psi \rightarrow \mu$ and $\Psi'_\perp \rightarrow \phi = e_0$. Then the optimum number r of iterations for Grover's original algorithm and for Grover's Microscope are equal. If we don't know ahead of time the value of $|\phi_p|$, then setting $\theta = \alpha$ will make both r and α depend on the unknown $|\phi_p|$, although the product $r\alpha$ will still be independent of it.

Let

$$\begin{aligned}
U_{\mu scope} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\
&= -e_1 e_0^T + e_0 e_1^T \\
&= -\Psi \Psi_{\perp}^T + \Psi_{\perp} \Psi^T.
\end{aligned} \tag{144}$$

Note that

$$U_{\mu scope} \Psi = \Psi_{\perp}. \tag{145}$$

From the point of view of quantum compiling, Grover's Microscope approximates the $\pi/2$ rotation $U_{\mu scope}$ by the r -fold product of $-R_{\Psi} R_{\Psi'_{\perp}}$, where we assume that $-R_{\Psi} R_{\Psi'_{\perp}}$ can be shown to have a SEO of low (polynomial in N_B) complexity. (If such a low complexity SEO cannot be found, then it is pointless to divide $U_{\mu scope}$ into r iterations of $-R_{\Psi} R_{\Psi'_{\perp}}$, and we might be better off compiling $U_{\mu scope}$ all at once.)

A Appendix: Deterministic $pd(Bool|Bool^{N_B})$ matrices

In this Appendix, we will first define a special kind of probability matrices which we call deterministic $pd(Bool|Bool^{N_B})$ matrices. Then we will show how such probability matrices can be q-embedded, and how their q-embedding can be expressed as a SEO.

Suppose $y \in Bool$ and $\vec{x} = (x^0, x^1, \dots, x^{N_B-1}) \in Bool^{N_B}$. Let $f : Bool^{N_B} \rightarrow Bool$.

We will say that f is **AND-like** if $f(\vec{x}) = \theta(\vec{x} = \vec{x}_{target})$ for some target vector $\vec{x}_{target} \in Bool^{N_B}$. An AND-like f maps all \vec{x} into zero except for \vec{x}_{target} which it maps into one. Thus, $|f^{-1}(1)| = 1$. An example of an AND-like f is the multiple AND gate $f(\vec{x}) = x^0 \wedge x^1 \wedge \dots \wedge x^{N_B-1}$, which can also be expressed as $f(\vec{x}) = \theta[\vec{x} = (1, 1, \dots, 1)]$.

We will say that f is **OR-like** if $f(\vec{x}) = \theta(\vec{x} \neq \vec{x}_{target})$ for some target vector $\vec{x}_{target} \in Bool^{N_B}$. An OR-like f maps all \vec{x} into one except for \vec{x}_{target} which it maps into zero. Thus, $|f^{-1}(0)| = 1$. An example of an OR-like f is the multiple OR gate $f(\vec{x}) = x^0 \vee x^1 \vee \dots \vee x^{N_B-1}$, which can also be expressed as $f(\vec{x}) = \theta[\vec{x} \neq (0, 0, \dots, 0)]$.

We will say that f has a **single target** if it is either AND-like or OR-like. If f has more than one target (i.e., if $|f^{-1}(0)|$ and $|f^{-1}(1)|$ are both greater than one), then we will say that f has **multiple targets**.

Suppose $y \in Bool$ and $\vec{x} = (x^0, x^1, \dots, x^{N_B-1}) \in Bool^{N_B}$. Let $f : Bool^{N_B} \rightarrow Bool$. In this section, we consider **deterministic $pd(Bool|Bool^{N_B})$ matrices**; that is, probability matrices of the form $P(y|\vec{x}) = \delta(y, f(\vec{x}))$. First let us consider the case that f has a single target. For example, for $N_B = 2$, if f is an AND gate

$$P(y|\vec{x})_{AND} = \left\{ \begin{array}{c|cccc} & & \begin{matrix} (x^0, x^1) \rightarrow \\ 00 \quad 01 \quad 10 \quad 11 \end{matrix} \\ y \downarrow \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix} & \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 & 1 \end{matrix} \end{array} \right\}, \quad (146)$$

and if f is an OR gate

$$P(y|\vec{x})_{OR} = \left\{ \begin{array}{c|cccc} & & \begin{matrix} (x^0, x^1) \rightarrow \\ 00 \quad 01 \quad 10 \quad 11 \end{matrix} \\ y \downarrow \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} & \begin{matrix} 0 & 1 \end{matrix} \end{array} \right\}. \quad (147)$$

Suppose bit value y is stored in the bit labelled τ . And suppose bit values $x^0, x^1, \dots, x^{N_B-1}$ are stored in the bits labelled $\vec{\kappa} = (\kappa_0, \kappa_1, \dots, \kappa_{N_B-1})$. Define e_j for all $j \in Z_{0, N_S-1}$ to be the N_S dimensional column vector with j th component equal to one and all other components equal to zero. Let $\Pi_j = e_j e_j^T$ and $\Pi_{targ} = \Pi_{j_{targ}}$, where $j_{targ} \in Z_{0, N_S-1}$ is the target state. Π_{targ} can be expressed as product of number operators. Indeed, if

$$j_{targ} = \sum_{i=0}^{N_B-1} x_{targ,i} 2^i, \quad (148)$$

then

$$\Pi_{targ} = \Pi_{j_{targ}} = \prod_{i=0}^{N_B-1} [n(\kappa_i) \theta(x_{targ,i} = 1) + \bar{n}(\kappa_i) \theta(x_{targ,i} = 0)]. \quad (149)$$

For example, if $j_{targ} = 0$ then $\Pi_{targ} = \bar{n}(\kappa_0) \bar{n}(\kappa_1) \dots \bar{n}(\kappa_{N_B-1})$.

An AND-like probability matrix $P(y|\vec{x})$ is q-embedded within the unitary matrix

$$U_{AND-like} = [A(y, \vec{x} | \tilde{y}, \vec{x})] = \begin{array}{c|cc} & \tilde{y} = 0 & \tilde{y} = 1 \\ y = 0 & 1 - \Pi_{targ} & -\Pi_{targ} \\ y = 1 & \Pi_{targ} & 1 - \Pi_{targ} \end{array}. \quad (150)$$

Note that

$$U_{AND-like} = 1 + \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} \otimes \Pi_{targ} \quad (151a)$$

$$= 1 + \Pi_{targ}(\vec{\kappa})(-i\sigma_y(\tau) - 1) \quad (151b)$$

$$= [-i\sigma_y(\tau)]^{\Pi_{targ}(\vec{\kappa})}. \quad (151c)$$

Eqs.(149) and (151c) show how to express $U_{AND-like}$ as a qubit rotation with multiple control qubits. Operations of this kind can be decomposed into a SEO using the techniques of Refs.[7] and [8].

An OR-like probability matrix $P(y|\vec{x})$ is q-embedded within the unitary matrix

$$U_{OR-like} = [A(y, \vec{x}|\tilde{y}, \vec{x})] = \frac{\begin{array}{c|cc} & \tilde{y} = 0 & \tilde{y} = 1 \\ \hline y = 0 & \Pi_{targ} & 1 - \Pi_{targ} \\ y = 1 & 1 - \Pi_{targ} & -\Pi_{targ} \end{array}}{\quad} . \quad (152)$$

Note that

$$U_{OR-like} = \begin{pmatrix} 0 & I_{N_S} \\ I_{N_S} & 0 \end{pmatrix} \begin{pmatrix} 1 - \Pi_{targ} & -\Pi_{targ} \\ \Pi_{targ} & 1 - \Pi_{targ} \end{pmatrix} \quad (153a)$$

$$= \sigma_x(\tau)[-i\sigma_y(\tau)]^{\Pi_{targ}(\vec{\kappa})} . \quad (153b)$$

Finally, let us consider the case when $f : Bool^{N_B} \rightarrow Bool$ has multiple targets. Let $T \subset Z_{0, N_S-1}$ be the set of these targets; i.e., either $T = f^{-1}(0)$ or $T = f^{-1}(1)$. Define Π_{targ} by

$$\Pi_{targ} = \sum_{j \in T} \Pi_j . \quad (154)$$

Π_{targ} can be expressed as a product of number operators. Indeed, each Π_j on the right hand side of Eq.(154) can be separately expressed, using Eq.(149), as a product of number operators. If $T = f^{-1}(1)$, then $P(y|\vec{x})$ is q-embedded within the unitary matrix

$$\begin{aligned} U_{multi-targ} &= [A(y, \vec{x}|\tilde{y}, \vec{x})] = \frac{\begin{array}{c|cc} & \tilde{y} = 0 & \tilde{y} = 1 \\ \hline y = 0 & 1 - \Pi_{targ} & -\Pi_{targ} \\ y = 1 & \Pi_{targ} & 1 - \Pi_{targ} \end{array}}{\quad} \\ &= [-i\sigma_y(\tau)]^{\Pi_{targ}(\vec{\kappa})} . \end{aligned} \quad (155)$$

B Appendix: Quasi-deterministic $pd(Bool|Bool^{N_B})$ matrices

In this Appendix, we will first define a special kind of probability matrices which we call quasi-deterministic $pd(Bool|Bool^{N_B})$ matrices. Then we will show how such probability matrices can be q-embedded, and how their q-embedding can be expressed as a SEO.

Suppose $y \in Bool$ and $\vec{x} = (x^0, x^1, \dots, x^{N_B-1}) \in Bool^{N_B}$. Let $f : Bool^{N_B} \rightarrow Bool$. In the previous appendix, we considered deterministic $pd(Bool|Bool^{N_B})$ matrices; that is, probability matrices of the form $P(y|\vec{x}) = \delta(y, f(\vec{x}))$. In this section, we will consider **quasi-deterministic $pd(Bool|Bool^{N_B})$ matrices**; that is, probability matrices of the form

$$P(y|\vec{x}) = \sum_{\vec{t}} \delta(y, f(\vec{t})) P(t^0|x^0) P(t^1|x^1) \dots P(t^{N_B-1}|x^{N_B-1}) , \quad (156)$$

where we sum over all $\vec{t} = (t^0, t^1, \dots, t^{N_B-1}) \in \text{Bool}^{N_B}$. Fig.19 shows a CB net representation of Eq.(156). Examples of quasi-deterministic $pd(\text{Bool}|\text{Bool}^{N_B})$ matrices are:

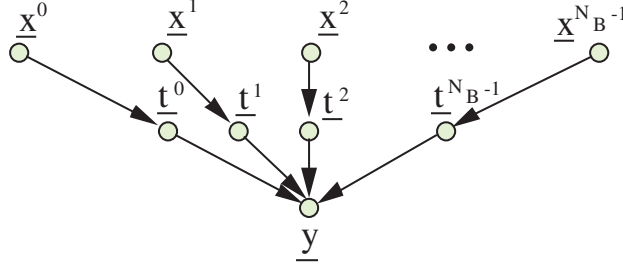


Figure 19: Quasi-deterministic (“noisy”) $pd(\text{Bool}|\text{Bool}^{N_B})$ gate.

(1)the **noisy OR**, for which $f(\vec{t}) = t^0 \vee t^1 \vee \dots \vee t^{N_B-1}$; (2)the **noisy AND**, for which $f(\vec{t}) = t^0 \wedge t^1 \wedge \dots \wedge t^{N_B-1}$; (3)the **noisy CNOT**, for which $f(\vec{t}) = t^0 \oplus t^1 \oplus \dots \oplus t^{N_B-1}$, etc.

For each $\alpha \in Z_{0, N_B-1}$, the probabilities $P(\underline{t}^\alpha = t | \underline{x}^\alpha = x)$ will be abbreviated by $p_{t,x}^\alpha$ for $t, x \in \text{Bool}$. $P(\underline{t}^\alpha = t | \underline{x}^\alpha = x)$ has two independent parameters which we may take to be p_{01}^α (the probability of false negatives) and p_{10}^α (the probability of false positives). p_{00}^α and p_{11}^α can be expressed in terms of these independent parameters: $p_{00}^\alpha = 1 - p_{10}^\alpha$, $p_{11}^\alpha = 1 - p_{01}^\alpha$. Whereas a completely general probability matrix $P(y|\vec{x}) \in pd(\text{Bool}|\text{Bool}^{N_B})$ has 2^{N_B} free parameters, a quasi-deterministic $P(y|\vec{x})$ has $2N_B$ free parameters.

Rather than q-embedding the probability matrix $P(y|\vec{x})$ as a whole, it is convenient to q-embed separately the probability matrices $P(y|\vec{t})$ and $P(t^\alpha|x^\alpha)$ for every $\alpha \in Z_{0, N_B-1}$. $P(y|\vec{t}) = \delta(y, f(\vec{t}))$ is a deterministic $pd(\text{Bool}|\text{Bool}^{N_B})$ matrix so its q-embedding is discussed in Appendix A. As for $P(t^\alpha|x^\alpha)$, it can be easily q-embedded as follows. For each $\alpha \in Z_{0, N_B-1}$, let

$$\Delta_p^\alpha = \begin{pmatrix} \sqrt{p_{00}^\alpha} & 0 \\ 0 & \sqrt{p_{01}^\alpha} \end{pmatrix}, \quad \Delta_q^\alpha = \begin{pmatrix} \sqrt{p_{10}^\alpha} & 0 \\ 0 & \sqrt{p_{11}^\alpha} \end{pmatrix}. \quad (157)$$

$P(t^\alpha|x^\alpha)$ is q-embedded within the unitary matrix:

$$[A(t^\alpha, \tilde{x}^\alpha | \tilde{t}^\alpha, x^\alpha)] = \begin{pmatrix} \Delta_p^\alpha & -\Delta_q^\alpha \\ \Delta_q^\alpha & \Delta_p^\alpha \end{pmatrix}. \quad (158)$$

Unitary matrices of this kind are called D-matrices in Ref.[8]. Ref.[8] shows how to decompose any D-matrix into a SEO.

References

- [1] R.R. Tucci, “Quantum Information Theory - A Quantum Bayesian Nets Perspective”, ArXiv eprint quant-ph/9909039 .
- [2] D. Deutsch and R. Jozsa, Proc. Roy. Soc. of London A (1992) **439**, 553. R. Jozsa, ArXiv eprint quant-ph/9707033 .
- [3] D.R. Simon, *Proceedings of the 35th Annual IEEE Symp. on the Found. of Comp. Sci.* (IEEE Computer Society, Los Alamitos, 1994). Extended Abstract on page 116. Full Version of the paper in S.I.A.M. Jour. on Computing, **26**, Oct 97.
- [4] E. Bernstein, U. Vazirani, Proceedings of the 25th Annual ACM Symposium on Theory of Computing, pages 11-20 (1993)
- [5] Lov K. Grover, ArXiv eprint quant-ph/9605043
- [6] T. Toffoli, *Automata, Languages and Programming, 7th Coll.* (Springer Verlag, 1980) pg. 632. E. Fredkin, T. Toffoli, Int. Jour. of Th. Phys. (1982) **21**, 219.
- [7] Barenco et al. “Elementary gates for quantum computation”, ArXiv eprint quant-ph/9503016
- [8] R.R. Tucci, “A Rudimentary Quantum Compiler(2cnd ed.)”, ArXiv eprint quant-ph/9902062 .
- [9] R.R. Tucci, “How to Compile a Quantum Bayesian Net”, ArXiv eprint quant-ph/9805016
- [10] B. Noble and J.W. Daniels, *Applied Linear Algebra*, Third Edition (Prentice Hall, 1988).
- [11] Grover’s algorithm expresses an orthogonal matrix as a product of real reflections. This is related to the QR decomposition of Linear Algebra[10], wherein any real (ditto, complex) matrix A is expressed as QR , where Q is a product of real (ditto, complex) “Householder” reflections and R is an upper triangular real (ditto, complex) matrix. A byproduct of the QR decomposition is a method for expanding an orthogonal (ditto, unitary) matrix as a product of real (ditto, complex) Householder reflections.
- [12] G. Brassard , P. Hoyer , M. Mosca , A. Tapp, ArXiv eprint quant-ph/0005055
- [13] Ahmed Younes, Jon Rowe, Julian Miller, ArXiv eprint quant-ph/0312022
- [14] P. Shor, *Proceedings of the 35th Annual IEEE Symp. on the Found. of Comp. Sci.* (IEEE Computer Society, Los Alamitos, 1994), page 124.

- [15] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W. Wootters, Phys. Rev. Lett., 70, 1895 (1993).
- [16] Note that the matrix A defined by Eq.(110) will have real entries if the $\xi^{(x)}$ basis is chosen to lie in the real $N_{\underline{x}}$ dimensional vector space and the Gram-Schmidt process is carried out in that same space. Thus, one can always find a q-embedding A for a probability matrix such that A is not merely unitary, but also orthogonal. However, if A is destined to become a node matrix in a QB net, it may be counterproductive to constrain A to be real, since this constraint may cause SEO decompositions of A to be longer.
- [17] S.L. Lauritzen and D.J. Spiegelhalter, Jour. of the Royal Statistical Society **B** (1988) **50**, 157.